

UNIVERSIDADE FEDERAL DE SANTA CATARINA
Centro de Ciências Físicas e Matemáticas – CFM

Uma formulação probabilística para o problema de síntese de populações estelares em galáxias *

João Rodrigo Souza Leão

Dissertação realizada sob orientação do Prof.
Dr. Roberto C. Fernandes Jr. e apresentada ao De-
partamento de Física da UFSC em preenchimento par-
cial dos requisitos para obtenção do título de Mestre
em Física.

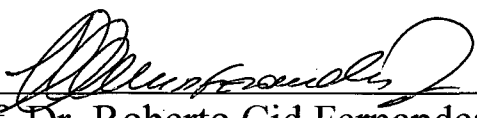
Florianópolis
2001

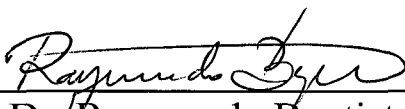
* Trabalho financiado pelo Conselho de Aperfeiçoamento de Pessoal do Ensino Superior (CAPES).

UMA FORMULAÇÃO PROBABILÍSTICA PARA O PROBLEMA DE SÍNTESE DE POPULAÇÕES ESTELARES EM GALÁXIAS


João Rodrigo Souza Leão

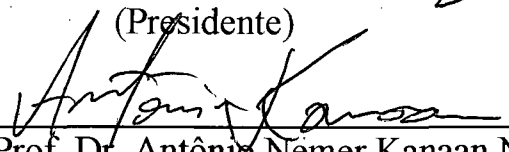
Esta dissertação foi julgada adequada para a obtenção do título de MESTRE EM FÍSICA, com especialização em ASTROFÍSICA, e aprovada em sua forma final pelo Programa de Pós-Graduação.


Prof. Dr. Roberto Cid Fernandes Jr.
(Orientador)

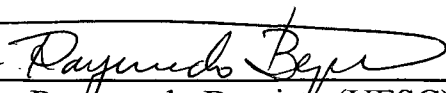

Prof. Dr. Raymundo Baptista
(Coordenador do curso)

BANCA EXAMINADORA


Prof. Dr. Roberto Cid Fernandes Jr.
(Presidente)


Prof. Dr. Antônio Nemer Kanaan Neto (UFSC)


Prof. Dr. Alex Schmidt (UFSM)


Prof. Dr. Raymundo Baptista (UFSC)
(Suplente)

Este trabalho é dedicado ao meu avô Nilberto Cavalcanti de Souza (*in memoriam*);
velho bonachão que cantando ou recitando, me ensinou a sonhar e a amar a poesia.

AGRADECIMENTOS

- À Deus por estar presente em todos os momentos.
- Aos meus pais por terem me dado a vida e a liberdade de seguir meus sonhos.
- À minha mãe por todo incentivo, amor e empenho em minha educação.
- Ao meu pai pela confiança e ensinamentos.
- À minha irmã (Luluzinha) pelo carinho e por baixar o som quando eu queria estudar.
- À todos os meus familiares e amigos pelo amor e incentivo.
- À minha avó Alba Lima de Souza pelo amor, incentivo e inestimável apoio em minha educação.
- Ao meu Tio Roberto Lima de Souza por ter me mostrado, desde muito cedo, a ciência e a natureza do pensamento científico.
- Ao meu tio Gilson Gomes de Medeiros por ter me dado (quando eu era ainda muito pequeno) as primeiras lições de Física.
- Aos meus tios e tias de Natal - RN: Alberto e Ana Lúcia, Roberto Lima e Socorro, Regina e Roberto Gurgel, Evaristo e Regane, Aurélio e Carminha, Lúcia e Edson, Lurdinha e Gilson, Eugênio e Catarina; por fazerem de minhas férias, todos os anos, tempos de muita alegria e boas lembranças.
- Aos meus primos (e os filhos dos meus primos), pelas verdadeiras lições de fraternidade que permearam nossa infância.
- A todos os meus professores (desde o tempo do jardim de infância) por se dedicarem à educação no Brasil.
- Ao meu orientador Roberto Cid Fernandes Jr. por sua paciência, dedicação, amizade e prazer em ensinar.
- Aos professores Raymundo Baptista e Antônio Nemer Kanaan do grupo de Astrofísica da UFSC por todo apoio, confiança e ensinamentos.
- Ao professor Alex Schmidt, da Universidade Federal de Santa Maria, pela enorme contribuição a este trabalho.
- Ao professor Claus Leitherer do STSci pela hospitalidade, amizade e ensinamentos valiosos.

- Aos colegas do Grupo de Astrofísica e do curso de Física da UFSC.
- Aos funcionários da Univesidade Federal de Santa Catarina pelo empenho e dedicação a esta unidade de ensino superior.
- À Sociedade Astronômica Brasileira e à Sociedade Astronômica Americana pelo apoio e promoção de eventos científicos.
- Ao Space Telescope Science Institute (STScI) pela oportunidade de participar do seu programa de verão de Maio a agosto de 2000.
- À CAPES pelo apoio financeiro e por acreditar na ciência brasileira.

Resumo

Utilizamos a **teoria de probabilidades** para abordar matematicamente o problema de sintetizar as propriedades espectrais de galáxias, usando uma base de populações estelares que varre intervalos de idade e metalicidade. Esta técnica fornece estimativas para o vetor de populações \vec{x} , que contém as contribuições dos diferentes elementos da base para o espectro integrado de uma galáxia. Também estimamos o avermelhamento A_V causado pela poeira inter-estelar. As estimativas são dadas na forma de distribuições de probabilidade, permitindo uma avaliação quantitativa das incertezas envolvidas. O formalismo é descrito no capítulo 2.

Desenvolvemos um código que sintetiza larguras equivalentes e cores (razões de fluxo) no contínuo a partir da combinação dos 12 elementos da base. Esta base é uma versão muito utilizada de uma base maior construída com as observações de aglomerados estelares feitas por Bica & Alloin (1986). Outros estudos com esta base forneceram estimativas médias para \vec{x} e A_V , usando o Procedimento das Combinações Diretas (PDC) para varrer o espaço de parâmetros. Neste trabalho, revisitamos este problema e introduzimos um algoritmo Markoviano quasi-randômico baseado no Método Monte Carlo de amostragem para varrer o espaço de parâmetros de 13-D e calcular a probabilidade $P(\vec{x}, A_V)$, guiados pela função Verossimilhança. Testes com este algoritmo são apresentados no capítulo 3.

No capítulo 4 analisamos a capacidade do procedimento de síntese de recuperar as histórias de formação estelar e enriquecimento químico como função dos (1) erros observacionais e (2) conjuntos de observáveis utilizados. Simulações foram realizadas para testar a habilidade do método de amostragem em mapear o espaço de parâmetros e encontramos que apenas para erros muito baixos ($S/N > 300$) todos os 12 componentes do vetor \vec{x} podem ser recuperados com precisão. A distribuição combinada em idade, entretanto, pode ser recuperada para erros tão baixos quanto 30. Concluimos que a base é capaz de recuperar com precisão a história de formação estelar, mas não é adequada para estudos sobre evolução química. Assim como outros estudos apontaram, verificamos que esta dificuldade é devida à estrutura interna da base. Concluimos que a síntese das larguras equivalentes em conjunto com as cores produz melhores resultados do que a síntese das larguras equivalentes apenas, apesar da inclusão de A_V como um parâmetro extra.

No capítulo 5 aplicamos o método a 24 das 55 galáxias do atlas de Kennicutt (1992). Estudamos a relação entre as populações estelares e os tipos morfológicos e espectrais das galáxias e uma excelente consistência geral entre estes diferentes tipos de informação foi encontrada.

No capítulo 6, ilustramos graficamente o “domínio sintético” gerado pelos 12 elementos da base definida por Schmidt *et al.* (1991). Este conceito (Pelat 1998) é muito útil para testar a possibilidade algébrica de sintetizar um dado grupo de observáveis e inferir sobre o problema da unicidade das soluções. Mapeamos 29 das 55 galáxias do atlas em várias projeções do domínio sintético e verificamos que nenhuma delas pode ser sintetizada exatamente pela base. De maneira geral, concluimos que a degenerescência estatística, devido aos erros observacionais, é um problema bem mais sério que a degenerescência algébrica, que surge quando se tenta sintetizar mais parâmetros que observáveis.

Finalmente, no capítulo 7 apresentamos as conclusões e indicamos as futuras aplicações da metodologia apresentada neste trabalho.

Abstract

Probability theory is used to provide a mathematical approach for the problem of synthesizing spectral properties of galaxies, using a base of simple stellar populations spanning a range of ages and metallicities. This technique, provides estimates for the population vector \vec{x} , containing the contributions of different base elements to the integrated spectrum of a galaxy, as well as the extinction A_V . The estimates are given in probability distributions, allowing a quantitative understanding of the uncertainties involved. The formalism is described in chapter 2.

A code was developed which synthesizes equivalent widths of absorption features and continuum colors from combinations of the 12 base elements of Schmidt *et al.* (1991). This base is a widely employed version of the much larger base built upon observations of star clusters by Bica & Alloin (1986). Most previous investigations with this base computed mean solutions for \vec{x} and A_V , using discrete combination procedures to sweep the parameter space. We revisit the problem introducing a Markov chain Monte Carlo sampling scheme to calculate the probability $P(\vec{x}, A_V)$ in a quasi-random walk through the 13-D parameter space, guided by the likelihood function. Tests of this scheme are presented in chapter 3.

In chapter 4 we analyze how well the synthesis procedure recovers the chemical enrichment and star formation histories as a function of (1) the observational errors and (2) the set of observables used. We tested the ability of the sampling scheme to map the parameter space by means of extensive simulations and found that only for extremely low errors ($S/N > 300$) *all* 12 base components can be recovered accurately. The age-binned distribution, however, is well recovered even for S/N as low as 30. We therefore conclude that the base reliably recovers the star formation history but it is not suitable for chemical evolution studies. As previous studies indicated, this difficulty is rooted on the internal structure of the base. We also found that synthesizing equivalent widths and colors produces better results than synthesizing only the equivalent widths, despite the inclusion of A_V as an extra parameter.

As an application of the method developed here, in chapter 5 we apply it to 24 of 55 galaxies observed by Kennicutt (1992). We also investigate the connection between the stellar populations and morphological and spectral types, finding an excellent overall consistency between these different types of information.

In chapter 6 we provide graphical illustrations of the “synthetic domain” spanned by the 12-element base. This concept, developed by Pelat (1998), is very useful to assess the algebraic synthetizability of a given set of observables and is thus useful to assess the problem of non-unique solutions. We plot 29 of the 55 galaxies of Kennicutt’s atlas in several projections of the synthetic domain and find that *none* of them can be exactly synthesized by the base! Overall, we find that the statistical degeneracy, induced by the presence of errors in the input data, is a more serious problem than the algebraic degeneracy which happens when one synthesizes more parameters than observables.

Finally, in chapter 7 we make our final remarks and indicate future applications of the method presented in this work.

SUMÁRIO

1. Introdução	11
1.1 O que é Síntese de Populações?	12
1.2 Técnicas de Síntese	12
1.2.1 Síntese Evolutiva (SPEE)	12
1.2.2 Síntese Semi-Empírica (SPESE)	13
1.3 Revisão dos trabalhos de SPESE	14
1.3.1 Os primeiros trabalhos	14
1.3.2 O Procedimento das Combinações Diretas (PCD), Bica (1988)	14
1.3.3 O Procedimento de Multi-Minimização (PMM), Schmidt <i>et al.</i> (1989, 1991)	15
1.3.4 A inclusão das cores no conjunto de observáveis, Schmitt <i>et al.</i> (1996)	17
1.3.5 O trabalho de Pelat (1997 e 1998)	17
1.4 Este trabalho	18
2. Formalismo	20
2.1 Equações básicas, vínculos e definições	20
2.2 Descrição da base e seus observáveis	22
2.3 Formulação probabilística do problema	25
2.3.1 A formulação Bayesiana da teoria da probabilidade: O teorema de Bayes	25
2.3.2 Aplicação do Teorema de Bayes ao problema de SPESE	27
2.4 O tratamento dos erros	29
2.5 As técnicas de amostragem: PCD e Metropolis	31
2.5.1 Amostragem Uniforme: PCD	32
2.5.2 Amostragem por Importância: O algoritmo de Metropolis	32
3. Testes do algoritmo de Metropolis	34
3.1 Relação entre N_{sim} e ϵ no processo Markoviano	34
3.2 Os posteriores projetados	35
3.3 A evolução do algoritmo	40
3.4 O limite da “temperatura infinita”	43
4. Testes da base com o algoritmo de Metropolis	46
4.1 O efeito da qualidade do espectro: A relação S/N	46
4.2 O efeito dos conjuntos de observáveis	54
4.3 Resultados agrupados em idade	54

4.4	A degenerescência entre idade e metalicidade	59
4.5	Resultados agrupados em metalicidade	59
4.6	Sumário dos resultados	64
5.	<i>Aplicação às galáxias do atlas de Kennicutt</i>	65
5.1	Os dados	65
5.2	Resultados	66
5.3	Conclusões	76
6.	<i>Domínio Sintético</i>	80
6.1	O conceito de Domínio Sintético	80
6.2	Visualização do Domínio sintético para a base de Schmidt <i>et al.</i> (1991)	81
6.2.1	O Domínio Sintético em diagramas W_i vs. W_i	81
6.2.2	O domínio sintético em diagramas Cor-Cor	86
6.3	Aplicação às galáxias do atlas de Kennicutt	86
7.	<i>Conclusões</i>	90
7.1	Conclusões finais	90
7.2	Aplicações futuras e perspectivas de continuidade	92
8.	<i>Apêndice</i>	94

LISTA DE FIGURAS

2.1	Variação das larguras equivalentes W_i em função do tempo. Cada cor refere-se a uma metalicidade diferente: 0.6 em rosa; 0.0, azul; -1.0 em verde e -2.0 vermelho. As linhas correspondentes estão indicadas no canto superior esquerdo em cada quadro. É clara a degenerescência entre idade e metalicidade: linhas metálicas fracas podem ser devidas a pouca idade ou baixa metalicidade Z	23
2.2	Variação das cores C_i em função do tempo. A convenção de cores é a mesma da figura anterior. Os comprimentos de onda correspondentes estão indicados à esquerda de cada quadro. Para os quatro painéis inferiores ($\lambda < 5870$) azul é para cima e vermelho para baixo, ao contrário dos painéis superiores ($\lambda > 5870$). Também aqui é também clara a degenerescência entre idade e metalicidade.	24
3.1	Distribuições de probabilidade para as componentes #1, #2 e #8 da galáxia 5, que respondem por 75% de sua luz. Estão representados 3 valores de N_{sim} : curva em vermelho (pontilhada)= 10^6 , verde (tracejada)= 10^7 e azul (contínua)= 10^8 . Horizontalmente podem ser lidos os valores de ϵ correspondentes.	37
3.2	Distribuições de probabilidade para a galáxia 15. Estão representadas as componentes #2, #9 e #11, que respondem por 70% da luz deste objeto. As convenções de cores são as mesmas da figura anterior. . . .	38
3.3	Curvas de distribuição de probabilidade para a galáxia 20. Estão representadas as componentes #2, #3 e #4, que respondem por 80% da luz deste objeto. As convenções de cores são as mesmas das 2 figuras anteriores.	39
3.4	Distribuição de probabilidades para todas as componentes e o avermelhamento para as galáxias 1 (em rosa,contínua) e 16 (azul, pontilhada). A qualidade do espectro e o conjunto de observáveis utilizado estão indicados	41
3.5	Ilustração da caminhada do algoritmo de Metropolis pelo espaço de soluções para a galáxia-fictícia 2. Os valores de entrada estão indicados por traços vermelhos na extremidade direita de cada painel. O primeiro painel indica a evolução do χ^2 para cada estado visitado, os dois painéis seguintes referem-se aos observáveis indicados e os seis painéis inferiores indicam a evolução dos x_i . As linhas pontilhadas no 2º e 3º painéis indicam os desvios $\sigma = \pm 1$ em torno do valor de entrada. Foi utilizado o conjunto A (todos os observáveis da base) e $S/N = 60$	42

3.6	Curvas de distribuição para as componentes 2 (azul), 6 (verde) e 9 (vermelho) da galáxia 4. Notar como os posteriores tendem para o valor teórico $\bar{x}_i = 8.33\%$, indicado pela curva em preto, no limite da temperatura infinita. As setas indicam os valores de entrada das componentes.	45
4.1	Curvas de distribuição para a galáxia-fictícia 3 (com população intermediária dominante) mostrando o efeito dos erros observacionais nos posteriores do vetor \vec{x} e A_V . O conjunto de observáveis está indicado e utilizamos 3 valores de S/N : 1000 (linha azul contínua), 100 (linha verde tracejada) e 10 (linha vermelha pontilhada). Vemos que os posteriores das componentes mais importantes se redistribuem entre outras componentes a medida que o ruído aumenta.	48
4.2	Distribuições de probabilidade para a galáxia-fictícia 21, onde dominam populações jovem, mostrando o efeito dos erros no vetor \vec{x} e A_V . O conjunto de observáveis está indicado e utilizamos 3 valores de S/N , seguindo a mesma convenção de cores da figura acima. O mesmo efeito de distribuição da figura anterior está presente.	49
4.3	O mesmo da figura anterior, inclusive as convenções de cores, porém para a galáxia-fictícia 6 (relativamente velha). Notar a redistribuição das componentes mais velhas a medida que cai a qualidade do espectro	50
4.4	Matriz de correlação para as componentes da galáxia-fictícia 3 (figura 4.1). A qualidade do espectro e o conjunto de observáveis estão indicados. Em cada painel x_i x x_j estão representados apenas 1 em cada 10^4 passos de uma simulação com $N_{sim} = 10^7$. Os valores reais estão indicados na parte inferior de cada coluna e também pelos círculos vermelhos. Podemos notar a forte anti-correlação entre elementos de mesma idade: x_1 e x_4 , x_5 e x_7 , x_{10} e x_{11}	51
4.5	Síntese de 12 galáxias fictícias (B1, B2,...,B12) onde apenas uma das componentes responde por 100% da sua luz. Isto equivale a sintetizar os elementos da base. No eixo vertical estão indicados os valores de x_{medio} e no outro eixo estão indicados os elementos da base. Três relações sinal/ruído estão representadas: 100 (branco), 30 (preto) e 10 (verde). Idealmente esperamos que os histogramas estivessem concentrados no observável que o gerou. No entanto vemos que alguns elementos são sintetizados por combinações de outros elementos da base. Isto ocorre principalmente com as componentes 2, 3 e 6.	52
4.6	Valores de entrada contra valores médios, para 100 galáxias-fictícias. Estão presentes duas relações S/N : círculos rosas indicam $S/N = 1000$ e as cruzes azuis $S/N = 60$. Notamos que mesmo para espectros bons ($S/N \sim 60$), é difícil recuperar as populações estelares.	53
4.7	Valores de entrada contra sintéticos (x_{medio}) para as 100 galáxias-fictícias e as combinações de conjunto de observáveis e S/N indicadas. Apenas alguns observáveis estão mostrados, mas a concordância é igualmente boa para todos os outros. Na última coluna esta representado o melhor modelo encontrado durante a simulação.	55

4.8	Resultados para a galáxia-fictícia 3 (com população intermediária dominante) mostrando o efeito dos conjuntos de observáveis nos posteriores do vetor \vec{x} e A_V . O valor de S/N está indicado e utilizamos 3 conjuntos de observáveis: A (linha azul contínua), B (linha verde tracejada) e C (linha vermelha pontilhada). Podemos notar como as distribuições ficam menos concentradas para grupos com menos observáveis.	56
4.9	Resultados para a galáxia-fictícia 21 (com população relativamente jovem) mostrando o efeito dos conjuntos de observáveis nos posteriores do vetor \vec{x} e A_V . O valor de S/N está indicado e utilizamos 3 conjuntos de observáveis, seguindo a mesma convenção da figura anterior.	57
4.10	O mesmo das figuras anteriores, mas a para a galáxia-fictícia 6 (com população relativamente velha). Seguimos as mesmas convenções de cores das 2 figuras anteriores.	58
4.11	Resultados agrupados em idade (mostradas à direita dos painéis) para o conjunto A e as relações sinal/ruído indicadas. Notar o enorme ganho em precisão, mesmo para espectros realísticos ($S/N \sim 30$). Comparar com a figura 4.6 que utiliza a mesma qualidade de espectro.	60
4.12	O mesmo da figura anterior mas para os conjuntos B e C. Os círculos correspondem a $S/N = 100$ e as cruzeiras a $S/N = 30$	61
4.13	Diferença de idades contra diferença de metalicidades (em escala logarítmica). Em cada painel estão indicados o conjunto de observáveis e a qualidade dos dados. O espalhamento é maior para espectros piores e para conjuntos de observáveis mais pobres. Notar como (na média) as idades são super-estimadas são compensadas por metalicidades sub-estimadas (e vice-versa).	62
4.14	Resultados combinados em metalicidade. Ao contrário do que se espera para os resultados combinados em idade, não se obtém resultados muito melhores.	63
5.1	Espectros integrados de três galáxias do atlas de Kennicutt. Os objetos estão ordenados a partir do topo em ordem decrescente de contribuição de populações velhas para sua luz.	68
5.2	Posteriores combinados em idade para a galáxia NGC3379 (elíptica). Como deveríamos esperar, vemos uma forte contribuição de populações velhas para a luz deste objeto.	72
5.3	Posteriores combinados em idade para a galáxia elíptica NGC4262. Notar que em relação à galáxia anterior existe uma contribuição mais forte da componente de idade $t = 10^9$ anos.	73
5.4	Posteriores combinados em idade para a galáxia NGC1832. Parte da luz desta galáxia provém de populações jovens e de regiões HII, como esperaríamos para uma galáxia espiral.	74
5.5	Posteriores para a galáxia NGC 4449 (Sm/Im). Boa parte da luz deste objeto provém de populações jovens e de regiões HII, certamente devido ao seu disco.	75

5.6	Tipo morfológico T em função da idade. Notar que a medida que o valor de T cresce, aumenta a contribuições das populações jovens para a luz dos objetos.	77
5.7	O mesmo da figura anterior, mas para as idades combinadas em apenas duas: $x_{velha} = 10^9$ e 10^{10} anos e $x_{nova} = 10^8, 10^7$ e 10^6 anos. . . .	78
5.8	Variação do tipo espectral em função da idade para as componentes x_{velha} e x_{nova} . Vemos que o crescimento do tipo espectral indica o aumento da contribuição das populações jovens para a luz do objeto. . . .	79
6.1	Projeção do domínio sintético no plano definido pelas W 's das linhas CaII K e MgI. Os números indicam a localização dos 12 elementos da base. As linhas conectando cada par de elementos indicam as W 's que resultam de combinações apenas deste par de elementos. Galáxias com W 's fora da envoltória definida por este emaranhado de linhas não podem ser modeladas pela base, ou seja, não existe nenhuma combinação fisicamente aceitável de elementos da base que reproduza de forma exata as W 's medidas.	83
6.2	Como a figura anterior, para outras combinações de linhas.	84
6.3	Como a figura anterior, para outras combinações de linhas.	85
6.4	Projeções do domínio sintético no plano definido por fluxos do contínuo, todos normalizados a 5870 Å. As curvas foram calculadas ignorando o efeito de avermelhamento.	87
6.5	Como a figura anterior, para outras combinações de cores.	88
6.6	Como a figura anterior, para outras combinações de linhas.	89

LISTA DE TABELAS

2.1	O números nesta tabela correspondem às componentes do vetor de populações $\vec{x} = (x_1, x_2, \dots, x_{12})$. Esta numeração será seguida ao longo de todo este trabalho.	25
2.2	Observáveis da base. Os fluxos no contínuo estão normalizados a 5870 Å.	26
3.1	As 26 galáxias-fictícias apresentadas em Schmidt <i>et al.</i> (1991). Horizontalmente podem ser lidos o número das galáxias e as frações do fluxo total (normalizadas em $\lambda 5870$ Å, representadas em porcentagem) das suas respectivas componentes. Os números das componentes correspondem àqueles indicados na tabela 2.1.	36
5.1	Na segunda coluna, vemos a classificação morfológica T, na terceira o Tipo espectral, definido por Sodr� & Cuevas (1997) e na quarta a classificação morfol�gica usual de Hubble.	66
5.2	Observ�veis calculados por H. Schmitt. Os fluxos est�o normalizadas em 5870 Å.	67
5.3	Resultados da s�ntese agrupados em idade para as 24 gal�xias do atlas de Kennicutt. Os erros acompanham as respectivas abund�ncias em %.	69
5.4	Observ�veis S�nteticos das 24 gal�xias do atlas de Kennicutt. Comparar esta tabela com os dados constantes na tabela 5.2.	70

Capítulo 1

Introdução à Síntese de Populações Estelares

O estudo da formação e evolução de galáxias é um dos problemas da Astrofísica atual que pode nos ajudar a responder muitas questões a respeito do universo e sua história. Entre elas podemos citar:

- Qual origem da grande variedade de tipos morfológicos de galáxias?
- Como entender os estágios iniciais da vida de uma galáxia?
- De que forma ocorreu o enriquecimento químico do universo?
- Como a interação gravitacional e os processos de colisão de galáxias contribuem para a formação de estrelas?

As respostas a estas perguntas, que motivam a Física e Astrofísica atuais, dependem de um entendimento preciso de como as galáxias se formam e evoluem.

Os parâmetros importantes para se estudar a história de formação e evolução de galáxias são certamente a **idade** e a **metalidade** das suas estrelas constituintes. Infelizmente, apenas para as galáxias mais próximas é possível observar cada estrela individualmente (Aloisi e Leitherer 2000). Ou seja: na maioria dos casos, dispomos apenas das propriedades espectrais **integradas**, de forma que não temos muitas informações a respeito dos tipos de estrelas (suas idades (t) e metalidades (Z)) existentes em galáxias de diferentes tipos morfológicos. Como estes parâmetros são fundamentais, ficamos impossibilitados de estudar a forma exata da evolução das galáxias, e assim dizer quais seus estados passados e futuros mais prováveis.

Já que o caminho observacional nos limita, temos que buscar outras maneiras de resolver o importante problema de se conhecer as populações estelares em galáxias. Poderia-se tentar uma abordagem puramente teórica através de princípios físicos fundamentais. Para isso, teríamos que considerar parâmetros como massa, tamanho, abundância química, momento angular e temperatura da proto-galáxia, além de efeitos extras como interação entre galáxias vizinhas (colisões, interação gravitacional, canibalismo, etc.). Não é preciso pensar muito para se descobrir que tal abordagem do problema é dispendiosa e imprecisa, uma vez que o número de

variáveis envolvidas é muito grande. Além disso, não conhecemos muitos dos processos físicos envolvidos. Não sabemos por exemplo, como (nem quando e com que taxa) uma massa de gás de digamos $10^5 M_\odot$ se transforma em estrelas, mesmo que tenhamos informações a respeito da sua temperatura, momento angular, composição química, etc.

Na verdade não precisamos nos preocupar com tais formas de abordar esta questão, pois existem técnicas e algoritmos computacionais, que nos possibilitam atacar o problema. Tais técnicas são chamadas de uma forma geral de **Síntese de Populações Estelares**.

1.1 O que é Síntese de Populações?

São técnicas que tentam recuperar a história de formação e enriquecimento químico de galáxias a partir da comparação de modelos com observações.

A questão é saber como decompor o espectro observado de uma galáxia em componentes correspondentes a grupos espectrais, ou bases de estrelas que representam diferentes populações estelares com diferentes metalicidades e estágios evolutivos.

O problema pode ser resumido de maneira formal pela seguinte equação:

$$F_\lambda^{obs} = \sum_j x_j F_\lambda^{j*} \quad (1.1)$$

onde F_λ^{obs} é o espectro observado do objeto, F_λ^{j*} representa um certo tipo de estrela ou população estelar simples (como um aglomerado estelar) com características conhecidas. Os x_j geram um vetor \vec{x} que contém as proporções de cada uma destas populações estelares. Portanto, o vetor \vec{x} fornece a abundância relativa de estrelas de cada tipo existente na galáxia. Como conhecemos as características destes grupos de estrelas, podemos inferir através do vetor \vec{x} a história de formação estelar e evolução química da galáxia.

O processo de síntese trata de encontrar o vetor de populações que melhor reproduz o espectro observado. Veremos a seguir as principais técnicas de síntese.

1.2 Técnicas de Síntese

Existem basicamente duas formas de se fazer Síntese de Populações Estelares (SPE). A saber: A chamada **Síntese de Populações Estelares Evolutiva (SPEE)**, que leva em conta modelos de evolução de estrelas e suas atmosferas; e uma outra que emprega **bases** de espectros observados de aglomerados estelares, conhecida como **Síntese de populações Estelares Semi-Empírica (SPESE)**.

1.2.1 Síntese Evolutiva (SPEE)

Este método, introduzido por Tinsley (1972), utiliza espectros de estrelas combinados com modelos de evolução estelar. Os espectros utilizados podem ser teóricos ou observados. Como se trata de uma abordagem direta para se estimar as populações estelares, são feitas hipóteses sobre determinados parâmetros do objeto que

se deseja estudar: Função de Massa Inicial (FMI), história de formação estelar e evolução química. O espectro sintetizado é então comparado com o espectro do objeto real. Caso apresentem grandes diferenças, os parâmetros são reajustados.

O grande problema desta estratégia é que existem muitos elementos sobre os quais temos apenas hipóteses (FMI, por exemplo) e outros sobre os quais nossas informações são imprecisas ou incompletas (evolução estelar). Ainda assim, esta técnica é capaz de produzir excelentes resultados, razão pela qual foi e continua sendo utilizada para o estudo de diversos problemas: evolução de galáxias, aglomerados estelares, galáxias Wolf-Rayet, etc. (Ver Chiosi e Renzini 1986, Barbuy e Renzini 1991, Leitherer *et al.* 1996 e 2000). Os programas GISELE (Bruzual & Charlot 1993 e 1996), PEGASUS (Fioc & Rocca-Volmerange 1997) e Starburst99 (Leitherer *et al.* 1999) se encontram entre os mais utilizados em estudos de SPEE.

1.2.2 Síntese Semi-Empírica (SPESE)

Os espectros de galáxias carregam grande quantidade de informação a respeito de suas histórias evolutivas em suas linhas de absorção ou emissão. A inclinação do contínuo (cores) também carrega este tipo de informação, embora seja mais susceptível aos efeitos de avermelhamento devido à poeira interestelar. Diante disso, devemos ser capazes de extrair estas informações dos espectros através da medida de observáveis: larguras equivalentes de linhas de absorção (ou emissão) e cores (razões de fluxo). Idealmente deveríamos sintetizar todo o espectro de um objeto, mas na prática estes observáveis é que são utilizados na síntese. Como o ponto de partida para se inferir sobre as populações estelares são os dados observacionais, esta abordagem é considerada uma técnica de **inversão**.

Além dos observáveis, utilizam-se bases de aglomerados observados (com idades e metalicidades conhecidas) ou grupos de estrelas divididas em classes espectrais. Os elementos da base ou os grupos de estrelas são combinados em diferentes proporções (os x_j da equação 1.1) e calculam-se os observáveis sintetizados. Os observáveis sintetizados são comparados com os observáveis reais do objeto, buscando-se sempre a combinação que melhor representa os observáveis reais. A combinação que fornece o menor desvio entre valores reais e sintetizados é considerada solução. Alternativamente, uma média de combinações que fornecem resultados aceitáveis pode ser considerada a solução do problema. Desta forma, fica determinado o vetor de populações \vec{x} que contém as proporções dos elementos da base que constituem o objeto. É claro que a forma de tratar os erros e como as soluções são testadas variam em cada trabalho.

Esta técnica é também conhecida como **Síntese de Populações Estelares com um banco de dados**, em alusão à base. Os primeiros trabalhos utilizando esta técnica surgiram a partir das décadas de 70 e 80 com Spinrad e Taylor (1971), Faber (1972) e Bica (1988). Posteriormente seguiram Schmidt *et al.* (1989) e (1991), Schmitt *et al.* (1996) e Pelat (1998). **Neste trabalho estamos interessados particularmente nesta segunda técnica de SPE.** Veremos a seguir um sumário destes trabalhos.

1.3 Revisão dos trabalhos de SPESE

1.3.1 Os primeiros trabalhos

O problema de se estimar a população de estrelas em galáxias foi, desde o início, encarado como um problema matemático, de estimativa de parâmetros. Sempre soube-se que a utilização de algoritmos numéricos e computacionais seriam de grande utilidade na tarefa de sintetizar a população de estrelas num objeto em estudo. Na década de 60 e 70, vários autores utilizaram procedimentos de síntese baseados em algoritmos de tentativa e erro: Wood (1966), Moore (1968) e Spinrad e Taylor (1971). Nestes trabalhos, as **cores** dos objetos observados eram ajustadas à modelos de índice de cor sintéticos. Estes são sem dúvida trabalhos pioneiros, mas falhavam em muitos casos, pois frequentemente se estimavam abundâncias negativas de certos grupos de estrelas. Outras vezes, obtinham-se diagramas H-R que contrariavam o conhecimento teórico a respeito de evolução estelar. Para isto, basta que o método super-estime grupos estelares em detrimento aos grupos que teoricamente deveriam originá-los. Estes efeitos são indicações claras de que esta abordagem é limitada e imprecisa, pois certos vínculos físicos eram desrepeitados.

Ficou claro, a partir de então, que desenvolvimentos futuros das técnicas de SPESE deveriam levar em conta em seu formalismo os vínculos físicos, que refletem nosso conhecimento teórico a respeito do sistema. Aparentemente, o primeiro autor a destacar esta necessidade foi Faber (1972).

Em seu trabalho, Faber propõe um formalismo que faz uso dos vínculos de positividade das abundâncias dos grupos de estrelas, do número de estrelas (os x_j), do número de estrelas e o da relação M/L (massa/luminosidade) desde o início. Os vínculos, desta forma, não eram utilizados apenas para testar soluções. Uma outra diferença importante é o abandono da abordagem por tentativa e erro, e a adoção de métodos de minimização. Neste caso, a grandeza minimizada é uma função que leva em conta o número de estrelas de um dado grupo espectral, x_j ; a luminosidade das estrelas em cada grupo, l_j ; e a luminosidade total da galáxia, L . Em seu trabalho, Faber sintetiza fotometria e não espectros, como em alguns trabalhos subsequentes. Através de um algoritmo de minimização é possível a estimativa do vetor de populações \vec{x} . Além disso, são computados os resíduos, entre estimativas e valores observados, dando meios para a interpretação dos erros e limitações do método. Deve-se também à Faber a introdução de uma interpretação algébrica para o problema, de uma forma bastante semelhante àquela que apresentaremos no capítulo 2. Uma das dificuldades desta abordagem é que a minimização da referida função era “guiada” por inequações que limitavam os valores dos parâmetros, mas os limites superior e inferior (das inequações) eram introduzidos arbitrariamente.

1.3.2 O Procedimento das Combinações Diretas (PCD), Bica (1988)

As transições atômicas e moleculares das atmosferas estelares refletem suas condições físicas bem como suas composições químicas. É por essa razão que larguras equivalentes são bons traçadores de idade e metalicidade em populações estelares. Os espectros observados das galáxias estão repletos destas transições e portanto suas larguras equivalentes contém informação acerca de suas histórias evolutivas.

O algoritmo formulado por Bica utiliza uma base composta de larguras equivalentes (9 linhas) para 35 espectros observados de aglomerados estelares com diferentes idades e metalicidades (conhecidas). A característica mais interessante deste método é que tanto a Função de Massa Inicial (FMI) como as atmosferas estelares e evolução estelar estão implícitas nos espectros dos aglomerados, dispensando portanto várias hipóteses empregadas em SPEE.

Para se determinar as abundâncias químicas e a distribuição de idades para uma dada galáxia, o código testa todas as possíveis combinações dos elementos da base que reproduzem as larguras equivalentes. Este procedimento é chamado de Procedimento de Combinações Diretas (PCD).

O espectro em estudo e consequentemente os observáveis, por serem fruto de observações, estão sujeitos a erros. Estes erros podem, naturalmente, influenciar os resultados da síntese. Este problema foi abordado adotando as combinações que reproduzem as larguras equivalentes (W_i) do espectro em estudo, dentro de uma janela de erro de 10%, sendo descartada qualquer outra combinação que reproduz os observáveis com um resíduo maior. Num segundo momento, é feito um refinamento com erros de 5%, apenas para a família de soluções encontrada na primeira etapa. É considerada solução a **média aritmética** das soluções encontradas na segunda etapa. Este procedimento é capaz de produzir bons resultados, mas **não possui nenhuma fundamentação matemática rigorosa**.

Outra questão é que o código desenvolvido por Bica varre um espaço de 35 dimensões onde cada x_j varia entre 0 e 100%. Isto implica que é necessário usar uma resolução muito baixa (da ordem de $\Delta x_j = 10\%$), devido ao tempo computacional. Na verdade, mesmo para resoluções baixas é muito difícil varrer o espaço de parâmetros devido à grande dimensionalidade do problema, forçando o estabelecimento de hipóteses a respeito dos caminhos evolutivos no plano t vs. Z , cujas dimensionalidades são menores e computacionalmente viáveis. Estes caminhos foram definidos segundo modelos simples de evolução química, segundo os quais a metalicidade só aumenta, desprezando efeitos de interação gravitacional e acreção de gás. Desta forma, esta técnica só é aplicável para sistemas isolados. Entretanto, foi através deste método que se pode, **pela primeira vez, traçar a evolução química de galáxias**. A unicidade da solução é discutida e a validade das soluções é testada através da análise dos caminhos evolutivos apontados por elas. A degenerescência algébrica (diferentes combinações dos elementos da base podem produzir o mesmo resultado) constitui outra grande dificuldade.

Este foi um trabalho pioneiro, que deu origem ao método mais popular de SPESE, e por isso vamos considerá-lo como ponto de partida em nosso trabalho. Alguns dos problemas apontados no trabalho de Bica foram seriamente considerados em trabalhos subsequentes, conforme veremos a seguir.

1.3.3 O Procedimento de Multi-Minimização (PMM), Schmidt et al. (1989, 1991)

Para resolver as dificuldades apontadas em 1.3.2, pode-se adotar uma das seguintes estratégias:

- Introduzir um tratamento matemático mais rigoroso, a fim de reduzir as incertezas do método.

- Reduzir o número de elementos da base, reduzindo assim o número de dimensões do espaço de parâmetros;
- Incorporar mais observáveis de forma a reduzir a degenerescência algébrica;

Nos trabalhos de Schmidt et al., as duas primeiras estratégias foram postas em prática.

No primeiro trabalho (1989), foi adotado um algoritmo de minimização que calcula médias entre as melhores soluções encontradas. Schmidt *et al.* apontam que resultados estatísticos são melhores que soluções únicas para o problema, que por ser degenerado, apresenta uma família de soluções igualmente boas, dentro da margem de erros dos dados. Uma função que leva em conta o quadrado das diferenças entre os valores sintéticos e observados das larguras equivalentes é minimizada através de um algoritmo específico, chamado pelo autor de **Procedimento de Multi-Minimização (PMM)**. Ao invés de se computar uma única solução para o problema de minimização, uma série de soluções-tentativa são testadas. Cada uma começa com um vetor de populações inicial $\vec{x} = 0$, exceto para uma de suas componentes, que é mantida num valor fixo durante a minimização. Durante as tentativas, cada componente assume valores de 0.1 até 1.0, com variação de 0.1. Desta forma, como o vetor \vec{x} possui 35 componentes, a síntese de um único objeto consiste de 350 tentativas (tarefas de minimização). Uma tentativa extra, sem componentes fixas é também calculada, permitindo assim 351 soluções para cada objeto. São descartadas soluções cuja diferença entre valores observados e sintéticos é maior que uma janela de erros pré-definida. A janela de erros é ligeiramente maior que os erros observacionais. Num segundo momento, usando como critério o χ^2 dos resíduos, são descartadas outras soluções. A solução final para um dado objeto é a média entre as soluções restantes. Desta forma, introduziu-se uma maneira de avaliar estatisticamente os resultados da síntese.

No segundo trabalho (1991), uma análise de componentes principais apontou que muitos dos 35 componentes da base eram **linearmente dependentes**, estando portanto a base sobre-dimensionada. Foi possível portanto, a redução da base para 12 elementos, reduzindo assim o espaço gerado pelos elementos da base. Os algoritmos utilizados em conjunto com a nova base, por Schmidt *et al.* para sintetizar as larguras equivalentes foram PCD (utilizada por Bica (1988)) e PMM (utilizada por Schmidt *et al.* (1989)). Diversas melhorias na aplicação destes algoritmos foram introduzidas, principalmente no modo de gerar as soluções e nos vínculos impostos ao problema.

Para visualizar os resultados, gráficos das larguras equivalentes observadas (W_i^{obs} , eixo y) e sintetizadas (W_i^{sin} , eixo x) são de grande utilidade, pois no caso ideal de precisão total, espera-se uma reta perfeita: $W_i^{obs} = W_i^{sin}$. Nos diagramas obtidos por Schmidt *et al.* os pontos se distribuem em torno da reta ideal, mostrando que o método adotado apresenta pequenos desvios e bons resultados. Outro resultado importante deste trabalho é a constatação de que a base é representativa o suficiente para recuperar as idades, mas não é suficientemente resolvida para se inferir sobre as metalicidades dos objetos.

Este trabalho, juntamente com o de Bica (1988), constitui um marco nas técnicas de SPESE, pois foram introduzidas diversas melhorias numéricas e estatísticas,

servindo de referência para outros trabalhos na mesma linha, inclusive para esta dissertação.

1.3.4 A inclusão das cores no conjunto de observáveis, Schmitt *et al.* (1996)

Os observáveis resumem o espectro dos objetos em estudo, e nesse sentido é ideal incluir tantos observáveis quanto possível na síntese. Como vimos acima, Bica (1988) e Schmidt *et al.* (1989, 1991) utilizaram apenas larguras equivalentes em seus estudos. Um passo natural no desenvolvimento das técnicas de SPESE seria a introdução de informação sobre a forma do contínuo.

A inclusão das cores (ou razões de fluxo) foi feita por Schmitt *et al.* (1996) para o estudo da galáxia NGC 6240, originária da colisão entre duas galáxias espirais. As cores incluídas são (em Å): 3290, 3660, 4020, 4570, 6630, 7520, 8700; todas normalizadas em 5870 Å. Estes observáveis foram originalmente utilizados em Bica & Alloin (1986), Bica & Alloin (1987) e Bica *et al.* (1994). É claro que a inclusão das cores tornou necessária a modelagem do avermelhamento. Embora utilizem um conjunto maior de observáveis na síntese, e uma versão melhorada do algoritmo de Bica (1988), ainda utilizou-se a técnica de PCD para varrer o espaço de parâmetros. Pela distribuição das abundâncias no plano t vs. Z , foi possível construir um cenário para se entender o processo de fusão dos dois núcleos originais e também a ionização do gás.

Este trabalho representa um retrocesso histórico nas técnicas de SPESE, pois nenhuma das melhorias introduzidas por Schmidt *et al.* (1991) foi utilizada: incluiu-se as cores no conjunto de observáveis, e uma versão melhorada do algoritmo de Bica (1988) foi utilizada, mas ainda utilizou-se a técnica de PCD para varrer o espaço de parâmetros e a base original composta por 35 elementos.

Esta técnica foi popularizada da maneira como foi formulada neste trabalho, mas utilizando a base reduzida por Schmidt *et al.* (1991). Alguns destes trabalhos são citados na seção 1.4.

1.3.5 O trabalho de Pelat (1997 e 1998)

Os desenvolvimentos mais recentes em técnicas de SPESE são devidos a Pelat. Sua principal contribuição foi elucidar, em Pelat (1997) e principalmente Pelat (1998), o problema da **não unicidade**, ou **degenerescência algébrica** em SPESE. Este problema ocorre pelo simples fato de que, na maioria dos casos, o número de observáveis sintetizado é menor do que o número de elementos da base (n_*), de modo que o sistema algébrico resultante é indeterminado, contendo **mais variáveis que equações**. Em Bica (1988) por exemplo, 35 variáveis eram deduzidas a partir de não mais que 9 medidas! Isto conferiu certo ceticismo quanto a confiabilidade de técnicas de SPESE, baseado na crença arraigada de que com mais parâmetros que equações existe um número infinito de soluções e portanto o problema é insolúvel.

Com um tratamento matemático elegante baseado em conceitos de álgebra linear e convexa, Pelat demonstrou que é possível obter uma base ortonormal para o espaço de soluções através da determinação de um sub-espaço de soluções do problema de n_{obs} equações e $n_* > n_{obs}$ variáveis, e pela determinação da intersecção desse com os

sub-espacos definidos por vínculos físicos de normalização e positividade (ver seção 2.1). As infinitas soluções são assim **delimitadas** à uma região conhecida do espaço de parâmetros, dentro da qual as larguras equivalentes observadas são reproduzidas exatamente.

Pelat mostrou ainda que em certas situações, o sistema tem solução **única** apesar de ser algebricamente indeterminado. De fato, a maioria das 26 galáxias-fictícias definidas por Schmidt *et al.* (1991) em seus testes pode ser **exatamente** sintetizada mesmo com menos equações do que variáveis! Outra contribuição de Pelat foi a definição do conceito de **Domínio Sintético**, que discutiremos no capítulo 6.

1.4 O presente trabalho: Uma formulação probabilística para o problema de SPESE

Neste trabalho exploraremos a síntese de população tal como formulada por Bica (1988). Assim como naquele trabalho, usaremos propriedades integradas de aglomerados estelares, com uma base para análise de populações estelares em galáxias. Tal como vimos nas seções anteriores, os trabalhos de Schmidt *et al.* e Pelat representaram grandes avanços nesta abordagem do problema de SPESE. É verdade, porém, que a maior parte das aplicações desta técnica de síntese foram realizadas com o algoritmo originalmente proposto e desenvolvido por Bica, ou com a variante desenvolvida por Schmitt *et al.* (1996). Os trabalhos de Schmidt *et al.* (1989), Bica *et al.* (1990), Jablonka *et al.* (1990), Schmidt *et al.* (1990), Jablonka & Alloin (1995), De Mello *et al.* (1995), Bonatto *et al.* (1996, 1998, 1999 e 2000), Schmitt *et al.* (1999), Kong e Cheng (1999) e Raimann *et al.* (2000) foram todos desenvolvidos seguindo esta linha. A técnica de minimização de Schmidt *et al.* (1991), apesar de superior em muitos aspectos, não foi tão amplamente empregada, e a de Pelat (1998), por ser ainda muito recente só foi utilizada por Boisson *et al.* (2000).

No entanto, apesar de popular, os seguintes pontos permaneceram em aberto na formulação proposta por Bica:

- Não foi implementada de uma maneira matematicamente rigorosa;
- O método de amostragem utilizado era pouco eficiente
- Ainda não foi devidamente testada;

O objetivo deste trabalho é justamente propor melhorias para a técnica introduzida por Bica. Para tanto, propomos um tratamento matemático-estatístico que inclui naturalmente o efeito dos erros observacionais. Propomos a utilização de um método de amostragem mais eficiente, baseada no **Algoritmo de Metropolis**. Estes dois tópicos constituem o formalismo que adotamos neste trabalho e serão apresentados no capítulo 2. Na tentativa de testar o método de amostragem, realizamos no capítulo 3 alguns testes que comprovam a eficácia do método. No capítulo 4 utilizamos nosso novo método para testar a base de Bica. Em particular vamos mostrar quantitativamente como a qualidade dos dados influencia a síntese e como a inclusão das cores afeta os resultados.

Como uma aplicação direta da metodologia aqui proposta, sintetizamos no capítulo 5 24 das 55 galáxias do atlas de Kennicutt (1992).

No capítulo 6 mostramos como pode-se determinar o domínio sintético da base, ou seja: como as relações internas entre os elementos da base determinam a possibilidade ou não de se sintetizar determinados objetos. Em especial, mostramos como as galáxias do atlas de Kennicutt se encontram distribuídas no domínio sintético da base.

Finalmente, no capítulo 7 apresentamos as considerações finais e as conclusões. Apontamos também os caminhos possíveis para este método, bem como para as técnicas de SPESE.

Capítulo 2

O Formalismo probabilístico e os métodos de amostragem

Neste capítulo propomos um formalismo probabilístico e também um novo método de amostragem para o problema de SPESE.

Devido aos erros observacionais, que conferem incertezas aos resultados, podemos esperar que exista uma probabilidade de um vetor de populações ser uma possível solução para um dado conjunto de observáveis. Um tratamento probabilístico computa distribuições em torno do valor real, no lugar de resultados que fornecem apenas os valores dos x_j . É por isso que propomos um formalismo que desde o início considere os erros como vínculos importantes; ao invés de serem tratados apenas para testar possíveis soluções. Às soluções do tipo distribuição, devido à sua largura, está associada a probabilidade de que valores sob a curva de distribuição representem uma boa solução para o problema. Vemos desta forma, que um tratamento estatístico não é apenas útil, mas fundamental para uma correta interpretação dos resultados da síntese. Qualquer desenvolvimento seguinte das técnicas de SPESE deve certamente priorizar este fato.

Neste capítulo vamos apresentar as **equações básicas** para o tratamento das grandezas envolvidas, uma breve revisão da **formulação Bayesiana** da teoria de probabilidades e sua aplicação ao presente problema, o **tratamento de erros** e ainda as **técnicas de amostragem**: PCD e o algoritmo de Metropolis, que propomos como uma maneira mais eficaz de mapear o espaço (\vec{x}, A_V) . Além disso, apresentamos também a base de aglomerados de Schmidt (1991) e os seus observáveis. Estes tópicos, que serão discutidos a seguir, constituem o método de SPESE que utilizaremos neste trabalho.

2.1 Equações básicas, vínculos e definições

Vamos primeiramente definir o problema matemático. Devemos sintetizar um conjunto de n_{obs} observáveis, a partir de uma base de aglomerados com n_* elementos; sendo $n_{obs} = n_W + n_C$, onde n_W é o número de larguras equivalentes (W_i) e n_C é o número de cores (C_i). Os observáveis são sintetizados a partir do vetor de populações \vec{x} , cujas n_* componentes expressam as frações do fluxo total, normalizados em um comprimento de onda de referência (em nosso caso 5870 Å). A extinção A_V deve

também ser considerada como um outro parâmetro a ser determinado, que atua somente sobre as cores C_i .

A determinação do vetor de populações \vec{x} (e a sua incerteza) é o grande objetivo do método. Este vetor de n_* dimensões e o avermelhamento estão sujeitos, entretanto, a vínculos, que conferem um sentido Físico e real ao problema de se estimar a idade (t) e a metalicidade (Z) das estrelas constituintes de uma galáxia.

As equações abaixo constituem os vínculos que devem ser respeitados pelos resultados da síntese:

$$\sum_{j=1}^{n_*} x_j = 1 \quad ; \quad x_j \geq 0 \quad ; \quad A_V \geq 0 \quad (2.1)$$

A primeira equação apenas limita o fluxo total da luz (em 5870 Å) à unidade, ou seja: constitui um vínculo de normalização. A segunda e a terceira apenas restringem as soluções para valores positivos ou nulos (vínculos de positividade), de acordo com as definições de x_j e A_V . Nota-se que a condição de normalização efetivamente reduz o número de parâmetros à n_* , sendo $n_* - 1$ componentes de \vec{x} mais A_V .

Os valores sintéticos para os observáveis são dados em termos dos parâmetros \vec{x} e A_V por:

$$W_i^{sin} = \frac{\sum_{j=1}^{n_*} W_{ij}^* F_{ij}^* x_j}{\sum_{j=1}^{n_*} F_{ij}^* x_j} \quad ; \quad i = 1..n_W \quad (2.2)$$

$$C_i^{sin} = R_i(A_V) \sum_{j=1}^{n_*} C_{ij}^* x_j \quad ; \quad i = 1..n_C \quad (2.3)$$

onde W_{ij}^* , F_{ij}^* , C_{ij}^* e $R_i(A_V)$ são respectivamente: a largura equivalente i devido à componente j da base, o Fluxo no contínuo sobre a linha i devido à componente j , a contribuição ao contínuo i da componente j , e a lei de avermelhamento escolhida, avaliada no respectivo comprimento de onda do fluxo C_i . Estas equações estão sujeitas aos vínculos de normalização e positividade das equações 2.1.

Escolhemos uma lei de avermelhamento polinomial conforme descrito no trabalho de Cardelli *et al.* (1989). O fato da lei de avermelhamento $R_i(A_V)$ não possuir uma dependência explícita em j implica em $A_{V_j} = A_V$. Esta hipótese é forte e assume que toda a galáxia (e por consequência todas as suas populações estelares) estão sendo avermelhadas pela mesma coluna de poeira. É claro que um modelo mais realístico assumiria diferentes avermelhamentos para o bojo e para os braços da galáxia, por exemplo.

Colocado desta forma, o problema fica matematicamente bem definido, podendo ainda ser interpretado algebricamente: A base gera um espaço de n_* dimensões. Os observáveis, que são calculados através das equações 2.2 e 2.3, são pesados pelo vetor de populações \vec{x} . Este pode ser escrito como:

$$\vec{x} = (x_1, x_2, x_3, \dots, x_{n_*}) \quad (2.4)$$

A síntese fornece como resultado as distribuições de probabilidade para as componentes x_j do vetor \vec{x} . Uma dada galáxia pode desta forma ser representada por

um vetor neste espaço de n_* dimensões. A extinção A_V , que também caracteriza a galáxia-alvo, desempenha o papel de um outro eixo neste espaço de n_*+1 dimensões.

2.2 Descrição da base e seus observáveis

Antes de discutirmos o formalismo de erros e o algoritmo de Metropolis para varrer o espaço de parâmetros, vamos mostrar a base e os observáveis compreendidos por ela. Conforme mencionamos na seção 1.4 vamos utilizar como base de referencia a base de 12 elementos de Schmidt *et al.* (1991).

As idades compreendidas pela base são 5: 10^6 , 10^7 , 10^8 , 10^9 e 10^{10} anos, sendo que a população mais velha corresponde às populações do tipo aglomerado globular e a mais nova (HII) corresponde a aglomerados ainda com formação de estrelas. As metalicidades são 4: 0.6, 0.0 -1.0 e -2.0. Ver tabela 2.1.

Os observáveis desta base compreendem $n_W=9$ larguras equivalentes: CaII K $\lambda 3933$, CN $\lambda 4200$, banda-G $\lambda 4301$, MgI $\lambda 5175$, CaII $\lambda 8543$, CaII $\lambda 8662$, H δ $\lambda 4102$, H γ $\lambda 4340$ e H β $\lambda 4861$ Å; e $n_C=7$ contínuos, selecionados em comprimentos de onda pivô: 3290, 3660, 4020, 4510, 6630, 7520 e 8700 Å, todos normalizados em 5870 Å. Reproduzimos na tabela 2.2 os valores destes observáveis, que foram publicados originalmente nos trabalhos de Bica e Alloin (1986b,1987), Schmitt *et al.* (1996) e Bica *et al.* (1994).

Podemos ver nesta segunda tabela que se trata de um conjunto máximo de $n_W + n_C = 16$ observáveis para modelar 13 parâmetros: as 12 proporções e ainda o avermelhamento A_V .

Vamos ver agora como os observáveis compreendidos pela base variam em função da idade do aglomerado e de suas metalicidades. Isto nos ajudará a verificar quais observáveis são bons traçadores de idade e/ou metalicidade.

Na figura 2.1 plotamos a variação com o tempo de seis das larguras equivalentes representadas na base. Em cada quadro está indicada a linha correspondente. Cada cor representa uma metalicidade diferente: 0.6 em rosa; 0.0 (metalicidade solar), azul; -1.0, verde e -2.0 em vermelho. O primeiro efeito que podemos notar é que o valor das larguras equivalentes W_i das linhas metálicas aumenta com a metalicidade, como esperado. Para as linhas CaII K, CN, banda-G e MgI vemos que até 10^8 anos seus valores variam muito pouco, fazendo destas linhas boas indicadoras de idade apenas para $t > 10^8$ anos, quando os valores das linhas sofrem variações consideráveis para todas as 4 metalicidades representadas. O tripleto do CaII ($\lambda \sim 8500$) não é um bom indicador de idade, pois não sofre variações para nenhuma das metalicidades, a partir de $t = 10^7$ anos. No entanto, no intervalo $10^6 < t < 10^7$ anos, a linha sofre grande variação devido ao surgimento de super-gigantes vermelhas. A linha do hidrogênio H δ , assim como outras linhas da série de Balmer do HI, por ser uma linha não metálica, não fornece informações a respeito da metalicidade; como pode ser visto, as curvas se sobrepõem. No entanto é um bom indicador de idade, particularmente para populações pós-formação estelar, de idade $\sim 10^8$ anos. (Ver Gonzalez-Delgado *et al.* 1998).

Na figura 2.2 plotamos a variação de seis das cores compreendidas pela base, com o tempo. Seus comprimentos de onda correspondentes estão indicados em cada quadro e a convenção de cores para cada metalicidade é a mesma da figura anterior.

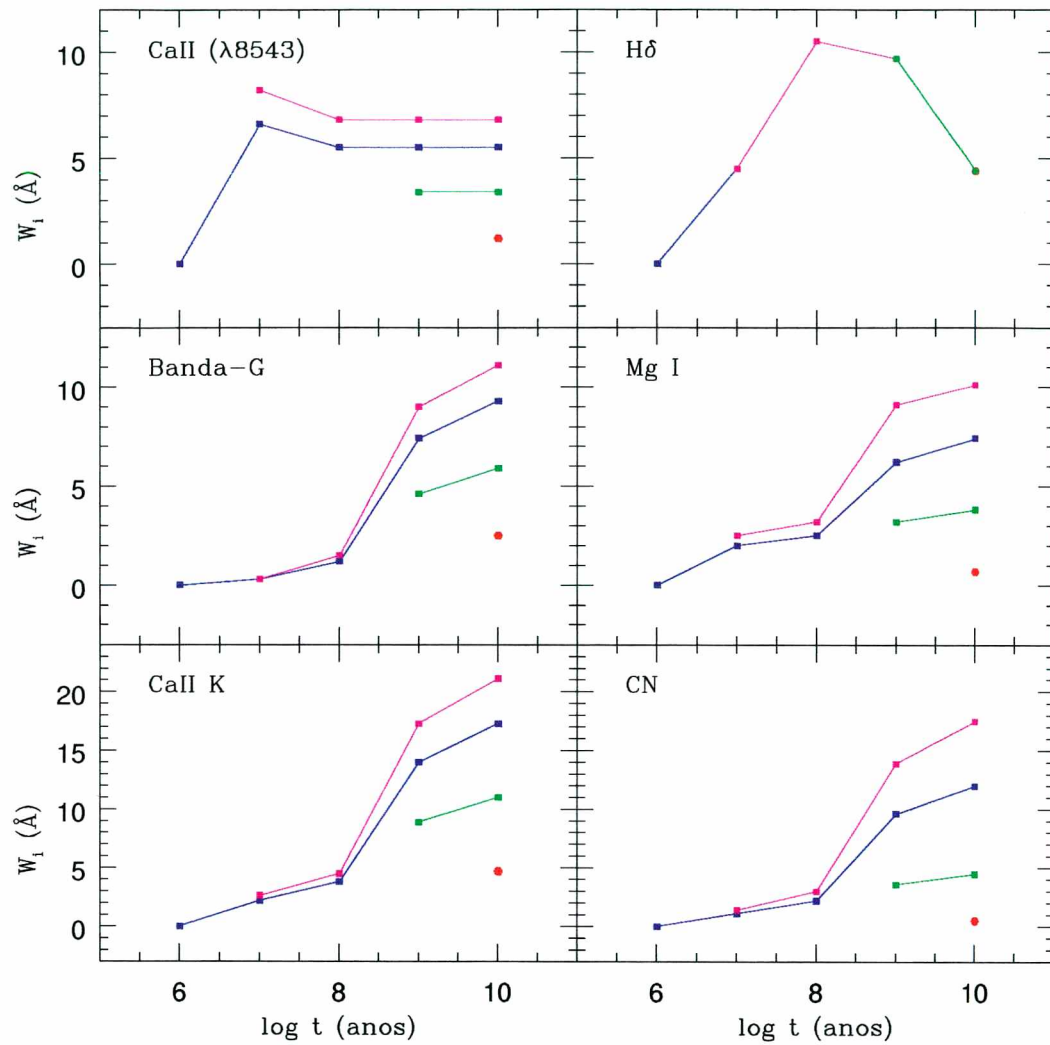


Fig. 2.1: Variação das larguras equivalentes W_i em função do tempo. Cada cor refere-se a uma metalicidade diferente: 0.6 em rosa; 0.0, azul; -1.0 em verde e -2.0 vermelho. As linhas correspondentes estão indicadas no canto superior esquerdo em cada quadro. É clara a degenerescência entre idade e metalicidade: linhas metálicas fracas podem ser devidas a pouca idade ou baixa metalicidade Z .

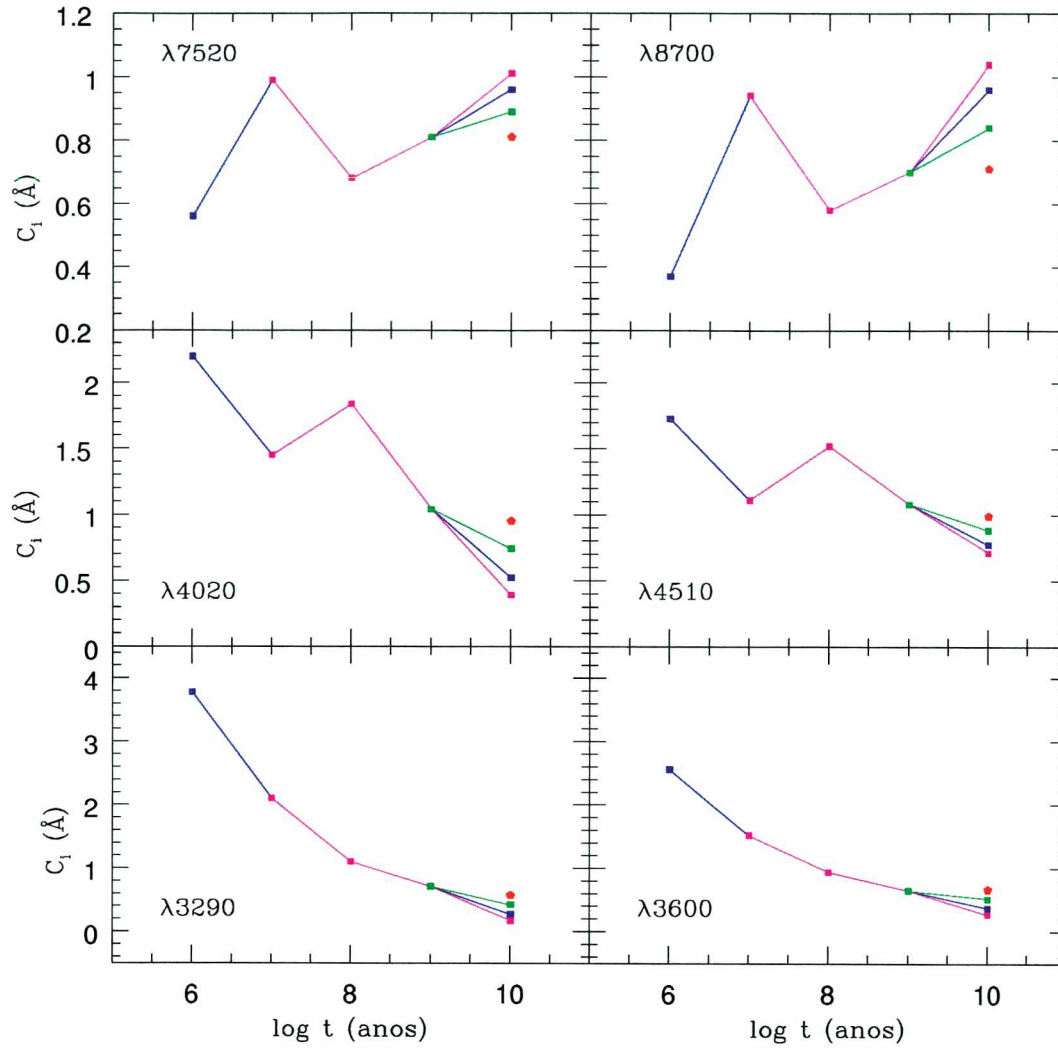


Fig. 2.2: Variação das cores C_i em função do tempo. A convenção de cores é a mesma da figura anterior. Os comprimentos de onda correspondentes estão indicados à esquerda de cada quadro. Para os quatro painéis inferiores ($\lambda < 5870$) azul é para cima e vermelho para baixo, ao contrário dos painéis superiores ($\lambda > 5870$). Também aqui é também clara a degenerescência entre idade e metalicidade.

Base Espectral de Schmidt <i>et al.</i> (1991)					
HII	10^7 anos	10^8 anos	10^9 anos	10^{10} anos	$\log [Z/Z_\odot]$
	# 10	# 8	# 5	# 1	0.6
#12	# 11	# 9	# 6	# 2	0.0
			# 7	# 3	-1.0
				# 4	-2.0

Tab. 2.1: O números nesta tabela correspondem às componentes do vetor de populações $\vec{x} = (x_1, x_2, \dots, x_{12})$. Esta numeração será seguida ao longo de todo este trabalho.

Primeiramente podemos observar que todos os contínuos são bons indicadores de idade, pois seus valores variam do azul para o vermelho com o tempo. No entanto, a figura mostra que as metalicidades se superpõem para todas as cores, exceto para $t > 10^9$ anos, só sendo possível obter informações sobre metalicidade no pequeno intervalo $10^9 < t < 10^{10}$ anos.

É clara a degenerescência entre idade e metalicidade, pois fica difícil inferir idades a partir de $t > 10^9$ anos. Não é claro se o espectro de um objeto está mais vermelho devido ao envelhecimento das suas populações ou devido ao aumento de sua metalicidade.

2.3 Formulação probabilística do problema

Neste trabalho, utilizamos elementos de uma formulação alternativa da teoria de probabilidades. Esta formulação recebe o nome de **Estatística Bayesiana** - em homenagem ao seu fundador, Thomas Bayes.

2.3.1 A formulação Bayesiana da teoria da probabilidade: O teorema de Bayes

A formulação Bayesiana se diferencia das outras abordagens porque seu formalismo é flexível o bastante para incorporar informações a respeito dos sistemas. Justamente por isso foi duramente criticada e tratada como subjetiva. Não pretendemos fazer uma revisão completa sobre o assunto, pois os trabalhos de Loredó (1990 e 1992) já expõem a Teoria Bayesiana de forma clara, inclusive para o leigo. Em seus trabalhos, Loredó salienta o poder e simplicidade desta formulação alternativa, concentrando nas inúmeras aplicações desta na Astrofísica. Assim, vamos expor aqui, apenas a demonstração do **Teorema de Bayes**, que será de grande utilidade neste trabalho.

Para chegarmos ao teorema de Bayes, podemos utilizar a regra do produto:

$$p(AB|C) = p(A|BC)p(B|C) = p(B|AC)p(A|C) \quad (2.5)$$

onde o termo da esquerda deve ser lido “probabilidade de A e B, dado C”. O mesmo vale para os outros termos. Vamos agora, mostrar uma aplicação simples mas bastante ilustrativa do uso desta formulação. Imaginemos que uma determinada hipótese H a respeito de um sistema precisa ser testada. Dispõe-se de um conjunto de dados D . Precisamos utilizar este conjunto de dados para calcular a plausibilidade da hipótese. Como dissemos, a formulação Bayesiana leva em seu formalismo

Larguras Equivalentes (Å)									
#	K	CN	G	MgI	CaT ₁	CaT ₂	H δ	H γ	H β
1	21.1	17.5	11.1	10.1	6.8	6.0	4.4	4.9	3.5
2	17.3	12.0	9.3	7.4	5.5	5.0	4.4	4.9	3.5
3	11.0	4.5	5.9	3.8	3.4	3.3	4.4	4.9	3.5
4	4.7	0.5	2.5	0.7	1.2	1.6	4.4	4.9	3.5
5	17.3	13.9	9.0	9.1	6.8	6.0	9.7	7.7	7.5
6	14.0	9.6	7.4	6.2	5.5	5.0	9.7	7.7	7.5
7	8.9	3.6	4.6	3.2	3.4	3.3	9.7	7.7	7.5
8	4.5	3.0	1.5	3.2	6.8	6.0	10.5	9.9	7.9
9	3.8	2.2	1.2	2.5	5.5	5.0	10.5	9.9	7.9
10	2.6	1.4	0.3	2.5	8.2	6.9	4.5	3.5	3.9
11	2.2	1.1	0.3	2.0	6.6	5.8	4.5	3.5	3.9
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Contínuo sobre as linhas									
#	K	CN	G	MgI	CaT ₁	CaT ₂	H δ	H γ	H β
1	0.34	0.48	0.55	0.87	1.03	1.04	0.44	0.58	0.79
2	0.48	0.59	0.65	0.90	0.96	0.96	0.56	0.67	0.83
3	0.72	0.78	0.81	0.95	0.84	0.84	0.76	0.82	0.91
4	0.94	0.96	0.97	1.01	0.72	0.71	0.96	0.97	1.00
5	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
6	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
7	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
8	1.89	1.75	1.68	1.23	0.59	0.58	1.79	1.65	1.38
9	1.89	1.75	1.68	1.23	0.59	0.58	1.79	1.65	1.38
10	1.55	1.34	1.27	1.06	0.95	0.94	1.40	1.25	1.08
11	1.55	1.34	1.27	1.06	0.95	0.94	1.40	1.25	1.08
12	2.34	2.03	1.93	1.37	0.40	0.38	2.11	1.90	1.56
Contínuo nos pontos pivô									
#	3290	3600	4020	4510	6630	7520	8700		
1	0.17	0.27	0.39	0.71	1.01	1.01	1.04		
2	0.27	0.37	0.52	0.77	0.99	0.96	0.96		
3	0.42	0.52	0.74	0.88	0.96	0.89	0.84		
4	0.57	0.67	0.95	0.99	0.92	0.81	0.71		
5	0.71	0.65	1.04	1.08	0.92	0.81	0.70		
6	0.71	0.65	1.04	1.08	0.92	0.81	0.70		
7	0.71	0.65	1.04	1.08	0.92	0.81	0.70		
8	1.10	0.94	1.84	1.52	0.82	0.68	0.58		
9	1.10	0.94	1.84	1.52	0.82	0.68	0.58		
10	2.10	1.52	1.45	1.11	0.95	0.99	0.94		
11	2.10	1.52	1.45	1.11	0.95	0.99	0.94		
12	3.78	2.56	2.20	1.73	0.74	0.56	0.37		

Tab. 2.2: Observáveis da base. Os fluxos no contínuo estão normalizados a 5870 Å.

outras informações relevantes ao problema em questão. Vamos chamá-las de I . Utilizando agora a equação 2.5, podemos escrever:

$$p(HD|I) = p(H|DI)p(D|I) = p(D|HI)p(H|I) \quad (2.6)$$

Podemos facilmente resolver a equação acima para o termo $p(H|DI)$, e obtemos:

$$p(H|DI) = p(H|I) \frac{p(D|HI)}{p(D|I)} \quad (2.7)$$

O termo da esquerda recebe o nome de **Posterior**. O termo $p(H|I)$ é chamado de **probabilidade a priori da hipótese H** . Esta expressão nos mostra como “atualizar” nosso conhecimento a respeito da hipótese, frente aos dados e as informações I (termo da esquerda), a partir do conhecimento a priori da probabilidade da hipótese frente às informações (termo $p(H|I)$). O outro termo é uma fração cujo numerador é a chamada **verossimilhança**, que mede a probabilidade $P(D|HI)$ dos dados ocorrerem, assumindo que a hipótese H esteja correta; e o denominador é a **verossimilhança global** (ou **evidência**), que pode ser considerada apenas como uma constante de normalização. Desta forma temos uma expressão analítica para verificar a plausabilidade da hipótese H , uma vez que se conheça o conjunto de dados D e as informações I .

A equação 2.7 é o **Teorema de Bayes**, cuja propriedade mais importante, é a sua capacidade de nos permitir aprender sobre as hipóteses a partir dos dados, que é justamente como funciona o processo científico.

2.3.2 Aplicação do Teorema de Bayes ao problema de SPESE

Para entendermos agora como aplicaremos esta formulação ao nosso problema, basta seguirmos a analogia:

- A hipótese H é o vetor de populações \vec{x} e a extinção A_V .
- O conjunto de dados D é justamente o conjunto de larguras equivalentes e cores que representam o espectro de uma galáxia. Os erros (σ) nas medidas destes observáveis podem ser considerados parte de D .
- I representa o conjunto de hipóteses envolvidas: representatividade da base, lei de avermelhamento, a natureza da extinção ($A_{V_i} = A_V$) e erros gaussianos.

Podemos agora facilmente reescrever a equação 2.7, utilizando as analogias acima:

$$p(\vec{x}, A_V|D, \sigma, I) = p(\vec{x}, A_V|\sigma, I) \frac{p(D|\vec{x}, A_V, \sigma, I)}{p(D|\sigma, I)} \quad (2.8)$$

Dispomos assim, de uma maneira de avaliar a probabilidade do par (\vec{x}, A_V) , em função dos dados, dos erros observacionais e de outras informações disponíveis.

Faremos agora a hipótese de que os erros observacionais obedecem a **distribuições Gaussianas**, com desvio padrão σ_i para cada observável (9 W_i e 7 C_i):

$$p(w_i) = A_i e^{-\chi_i^2} \quad ; \quad \chi_i^2 = \frac{w_i^{obs} - w_i^{sin}}{\sigma(w_i)} \quad (2.9)$$

onde A_i é uma constante.

Podemos assim, obter uma expressão para a verossimilhança total:

$$p(D|\vec{x}, A_V, \sigma, I) = \prod_{i=1}^{n_{obs}} A_i e^{-\chi_i^2/2} = \left[\prod_{i=1}^{n_{obs}} A_i \right] e^{-\sum_{i=1}^{n_{obs}} \chi_i^2/2} \quad (2.10)$$

ou seja: a verossimilhança é uma distribuição gaussiana em torno do valor mais provável para o observável i . O prior, na equação 2.8, não impõe quaisquer outros vínculos além daqueles da equação 2.1. Este tipo de prior é chamado de prior não informativo. Neste caso, a probabilidade posterior $p(\vec{x}, A_V|D, \sigma, I)$ é proporcional à verossimilhança. Utilizando as equações 2.8 e 2.10, temos:

$$p(\vec{x}, A_V|D, \sigma, I) \propto e^{-\chi^2(\vec{x}, A_V)/2} \quad (2.11)$$

Esta expressão representa o resultado do processo de síntese em nossa formulação.

Como temos duas classes de observáveis, o χ^2 se divide em uma parte associada à modelagem das larguras equivalentes e outra associada às cores, essa última dependente do avermelhamento A_V . Assim, podemos escrever:

$$\chi^2 = \chi_W^2(\vec{x}) + \chi_C^2(\vec{x}, A_V) = \sum_{i=1}^{n_W} \left(\frac{W_i^{obs} - W_i^{sin}(\vec{x})}{\sigma(W_i)} \right)^2 + \sum_{i=1}^{n_C} \left(\frac{C_i^{obs} - C_i^{sin}(\vec{x}, A_V)}{\sigma(C_i)} \right)^2 \quad (2.12)$$

onde: “sin” e “obs” denotam valores sintéticos e observados respectivamente. A letra σ representa o erro para o observável indicado.

Para obtermos as probabilidades posteriores para cada uma das componentes x_j do vetor de populações devemos **marginalizar** com respeito às demais componentes $x_{k \neq i}$:

$$p(x_j|D, \sigma, I) = \int \dots \int p(\vec{x}, A_V|D, \sigma, I) dx_1 \dots dx_{j-1} dx_{j+1} \dots dx_{n_*} dA_V \quad (2.13)$$

o mesmo pode ser feito para A_V . Essa marginalização simplesmente projeta o posterior (de 13 dimensões) $p(\vec{x}, A_V|D, \sigma, I)$ sobre o eixo x_j , produzindo a distribuição de probabilidade para a fração x_j do fluxo total devida à estrelas com idade e metalicidade j (ver Tabela 2.1).

Resolve-se assim o problema de se considerar os erros como um fator determinante dos resultados da síntese. Dentro deste formalismo, os erros observacionais das larguras equivalentes e das cores são levados em conta desde o início. Como resultado, obtemos distribuições de probabilidade para o par (\vec{x}, A_V) . O resultado, apresentado desta forma, nos permite avaliar de **forma quantitativa como a qualidade dos dados que dispomos influenciam os resultados**. Como consideramos os erros para cada observável em separado, poderemos avaliar também **como o conjunto de observáveis utilizado afeta a síntese**. Na verdade, estes

constituem dois objetivos específicos deste trabalho, como dissemos no capítulo 1, seção 1.4.

2.4 O tratamento dos erros

Sabemos que o espectro, e portanto os observáveis a serem sintetizados, estão sujeitos a incertezas devido aos erros observacionais. Para efeitos de testes do método precisamos além das galáxias fictícias, uma receita de erros a ser utilizada nas simulações adiante. Vamos começar pelas definições gerais:

F_λ = espectro observado,

C_λ = pseudo-contínuo,

$\Delta\lambda$ = largura da janela sobre a qual certa W_i é medida,

$\delta\lambda$ = resolução espectral,

λ_L = comprimento de onda do centro da janela.

A definição de largura equivalente é:

$$W = \int_{\Delta\lambda} \frac{C_\lambda - F_\lambda}{C_\lambda} d\lambda \sim \frac{\int_{\Delta\lambda} (C_\lambda - F_\lambda) d\lambda}{\bar{C}} \quad (2.14)$$

Definindo agora:

$$\bar{C} \equiv \int_{\Delta\lambda} \frac{C_\lambda d\lambda}{\Delta\lambda} ; \quad \bar{F} \equiv \int_{\Delta\lambda} \frac{F_\lambda d\lambda}{\Delta\lambda} \quad (2.15)$$

e usando as definições acima, W fica:

$$W = \frac{\bar{C}\Delta\lambda}{\bar{C}} - \frac{\bar{F}\Delta\lambda}{\bar{C}} = \Delta\lambda \left(1 - \frac{\bar{F}}{\bar{C}} \right) \quad (2.16)$$

Vemos então que $W = W(\bar{F}, \bar{C}, \Delta\lambda)$, sendo que \bar{F} e \bar{C} são fontes de incerteza em W. Para encontrarmos o erro em W devido a erros em \bar{F} e \bar{C} , vamos utilizar a fórmula de propagação de erros usual:

$$\epsilon^2(W) = \left(\frac{\partial W}{\partial \bar{F}} \right)^2 \epsilon^2(\bar{F}) + \left(\frac{\partial W}{\partial \bar{C}} \right)^2 \epsilon^2(\bar{C}) \quad (2.17)$$

Utilizando agora a equação 2.16 e calculando as derivadas parciais, temos:

$$\epsilon^2(W) = \left(\Delta\lambda \frac{\epsilon(\bar{F})}{\bar{C}} \right)^2 + (\Delta\lambda - W)^2 \left(\frac{\epsilon(\bar{C})}{\bar{C}} \right)^2 \quad (2.18)$$

vamos chamar os termos da direita respectivamente de ϵ_{WF}^2 e ϵ_{WC}^2 .

Precisamos agora trabalhar os termos $\epsilon(\bar{F})$ e $\epsilon(\bar{C})$; sendo que:

- \bar{F} é determinado por integração direta do espectro e $\epsilon^2(\bar{F})$ depende da qualidade (S/N) e do número \mathcal{N} de pontos na integração, i.e. da resolução espectral.
- \bar{C} é determinado por um julgamento arbitrário da localização do pseudo-contínuo sobre a janela. Seu valor também depende da qualidade dos dados, mas indiretamente, uma vez que sua posição é determinada de uma forma subjetiva (ver Cid Fernandes *et al.* 1998).

A equação 2.15, para \bar{F} , na prática é escrita como: escrita como:

$$\bar{F} = \frac{\sum_{i=1}^{\mathcal{N}} F(\lambda_i) \delta\lambda_i}{\Delta\lambda} = \frac{\delta\lambda}{\Delta\lambda} \sum_{i=1}^{\mathcal{N}} F_i \quad (2.19)$$

Podemos agora calcular $\epsilon^2(\bar{F})$:

$$\epsilon^2(\bar{F}) = \sum_{i=1}^{\mathcal{N}} \left(\frac{\partial \bar{F}}{\partial F_i} \right)^2 \epsilon^2(F_i) = \left(\frac{\delta\lambda}{\Delta\lambda} \right)^2 \sum_{i=1}^{\mathcal{N}} \epsilon^2(F_i) \quad (2.20)$$

e fazendo:

$$\sum_{i=1}^{\mathcal{N}} \epsilon^2(F_i) = \mathcal{N} \overline{\epsilon^2(F_i)} \quad (2.21)$$

podemos reescrever a equação 2.20 como:

$$\epsilon^2(\bar{F}) = \frac{\delta\lambda^2}{\Delta\lambda^2} \mathcal{N} \overline{\epsilon^2(F_i)} = \frac{\delta\lambda}{\Delta\lambda} \overline{\epsilon^2(F_i)} \quad (2.22)$$

onde usamos $\mathcal{N} \sim \Delta\lambda/\delta\lambda$.

Como temos agora uma expressão para o erro devido à \bar{F} , podemos reescrever o primeiro termo a direita (ϵ_{WF}) da equação 2.18, utilizando a equação anterior:

$$\epsilon_{WF}^2 = (\Delta\lambda)^2 \frac{\epsilon^2(\bar{F})}{\bar{C}^2} = \Delta\lambda \delta\lambda \frac{\overline{\epsilon^2(F_i)}}{\bar{C}^2} = \frac{\Delta\lambda \delta\lambda}{(S/N)_\lambda^2} \quad (2.23)$$

onde utilizamos o fato que :

$$\frac{\sqrt{\overline{\epsilon^2(F_i)}}}{\bar{C}} = "N/S" \quad (2.24)$$

é a razão ruído/sinal no contínuo adjacente à linha. Vamos introduzir uma notação e chamar o numerador da esquerda na equação acima simplesmente de ϵ_f .

Vamos agora explicitar um pouco mais $(S/N)_\lambda$:

$$(S/N)_\lambda = \frac{\bar{C}_\lambda}{\epsilon_{F,\lambda}} = \left(\frac{\bar{C}_\lambda}{\bar{C}_{5870}} \right) \left(\frac{\bar{C}_{5870}}{\epsilon_{f,5870}} \right) \left(\frac{\epsilon_{f,5870}}{\epsilon_{f,\lambda}} \right) \quad (2.25)$$

As três frações da direita na equação acima são respectivamente: o termo de cor, a razão sinal/ruído em 5870 Å $((S/N)_{5870})$ e a dependência em λ do ruído.

Escolhemos $\epsilon_{f,5870} = \epsilon_{f,\lambda}$, tal que $\epsilon_{f,\lambda}$ seja constante ao longo de todo o espectro. Desta forma, uma vez que o termo de cor é conhecido, o único parâmetro definindo $(S/N)_\lambda$ e portanto ϵ_{WF} é $(S/N)_{5870}$.

Assim, substituindo 2.25 em 2.23 temos:

$$\epsilon_{WF} = \frac{\sqrt{\Delta\lambda\delta\lambda}}{(S/N)_\lambda} = \frac{\sqrt{\Delta\lambda\delta\lambda}}{\left[\frac{\overline{C}_\lambda}{\overline{C}_{5870}}\right] \left[\left(\frac{S}{N}\right)_{5870}\right]} \quad (2.26)$$

Vamos agora reescrever o segundo termo da direita na equação 2.18 (ϵ_{WC}):

$$\epsilon_{WC} = (\Delta\lambda - W) \left[\frac{\epsilon(\overline{C})}{\overline{C}} \right]_{pc} \quad (2.27)$$

Por analogia com a equação 2.24, temos:

$$\left[\frac{\epsilon(\overline{C})}{\overline{C}} \right]_{pc} = \left(\frac{S}{N} \right)_{pc}^{-1} \quad (2.28)$$

As quantidades $(S/N)_\lambda$ e $(S/N)_{pc}$, (onde “pc” significa pseudo-contínuo), são proporcionais e a constante de proporcionalidade foi determinada em Cid Fernandes *et al.* (1998):

$$(S/N)_{pc} = \rho (S/N)_\lambda \quad ; \quad \rho \sim 3 \quad (2.29)$$

Substituindo 2.29 e 2.28 em 2.27, obtemos:

$$\epsilon_{WC} = \frac{(\Delta\lambda - W)}{3 \left(\frac{S}{N} \right)_\lambda} \quad (2.30)$$

Assim, a receita final pode ser escrita como:

$$\epsilon_W^2 = \epsilon_{WF}^2 + \epsilon_{WC}^2 \quad (2.31)$$

onde ϵ_{WF} é dado pela equação 2.26 e ϵ_{WC} é dado pelas equações 2.30.

2.5 As técnicas de amostragem: PCD e Metropolis

Na seção 2.3 vimos que o formalismo nos conduz à árdua tarefa de calcular integrais num espaço de 12 dimensões.

Conforme discutimos nas seções 1.3.2, 1.3.3 e 1.3.4, os referidos trabalhos utilizam uma forma de varredura do espaço de parâmetros denominada PCD, que constitui uma discretização do espaço. Por razões que discutiremos nas seções 2.5.1 e 2.5.2, abandonamos esta técnica de amostragem e adotamos uma outra baseada no **Método Monte Carlo de Amostragem por Importância** e no **algoritmo de Metropolis**. A seguir apresentaremos a técnica de PCD e suas maiores deficiências. Em seguida mostraremos como o Método Monte Carlo e algoritmos de amostragem por importância podem nos ajudar a sanar algumas dificuldades e varrer o espaço de parâmetros de uma maneira mais eficiente.

2.5.1 Amostragem Uniforme: PCD

A maneira mais simples e direta de se estimar a probabilidade de um certo par (\vec{x}, A_V) , utilizando a equação 2.13, é dividir o espaço (\vec{x}, A_V) em uma malha uniforme com resolução Δx para as frações x_i , e ΔA_V para a extinção. Note que a equação 2.13 para o posterior de x_i é simplesmente a **probabilidade média** $P(\mathbf{x}, A_V | D, \sigma, I)$ sobre o espaço gerado por todos os possíveis valores $x_{j \neq i}$ e A_V para um dado x_i . Assim, podemos aproximar a equação 2.13 pela soma:

$$P(x_i | D, \sigma, I) \simeq \frac{\sum_s e^{-\mathcal{E}(\mathbf{x}_s, A_{V,s})} \delta(x_i - x_{i,s})}{\sum_s e^{-\mathcal{E}(\mathbf{x}_s, A_{V,s})}} \quad (2.32)$$

onde s é um ponto (ou um “estado”) nesta malha de 13-Dimensões. O termo $\delta(x_i - x_{i,s})$, representando uma delta de Dirac, conserva apenas os pontos dessa malha nos quais a população i contribui com uma fração x_i de luz total da galáxia.

Esta forma de varrer o espaço de parâmetros é bastante fácil de ser implementada em um algoritmo, e foi a escolhida pela maioria dos trabalhos que utilizam a base de referência construída por Bica. A desvantagem mais notável deste procedimento é o tempo computacional necessário para se produzir resultados satisfatórios. Explica-se: para se obter distribuições de probabilidade “suaves” (boa amostragem do espaço), é preciso utilizar resoluções Δx e ΔA_V relativamente altas. O problema é que a resolução adotada interfere de uma maneira não linear no número de pontos da malha. Quanto maior a resolução, maior o número de pontos, e maior o tempo gasto para “visitá-los”. Por exemplo: para uma resolução $\Delta x = 10\%$ e $n_* = 12$, temos 3.5×10^5 pontos a serem visitados, aumentando para $\sim 8 \times 10^7$ se $\Delta x = 5\%$ e para mais de 4×10^{11} se utilizarmos uma resolução de 2% . O número total de pontos da malha é obtido multiplicando-se estes números pelo número de pontos da extinção definido por ΔA_V . Como vemos, o número de pontos cresce rapidamente conforme se melhora a resolução do método. Mesmo que fosse possível visitar cada um destes pontos, isto não seria muito útil: muitos deles contribuem pouco para a integral na equação 2.13, pois estão em regiões pouco prováveis do espaço de parâmetros (pequena verossimilhança).

2.5.2 Amostragem por Importância: O algoritmo de Metropolis

As técnicas de amostragem simples, possuem o inconveniente descrito acima: visitam-se, sem nenhum critério, todos os pontos do espaço sem levar-se em conta suas plausibilidades relativas. É neste contexto que se encaixam as **técnicas de amostragem por importância**, que varrem o espaço de uma maneira mais eficiente, mapeando a distribuição de probabilidade (equação 2.11). Um destes algoritmos é o de Metropolis *et al.* (1953), que é capaz de amostrar preferencialmente regiões onde o posterior é grande. O algoritmo de Metropolis, na forma como o utilizamos neste trabalho, pode ser descrito como segue:

1. Começamos de um ponto arbitrário do espaço (\vec{x}, A_V) ;
2. A cada iteração s , escolhe-se uma das 13 variáveis (12 x_j ou A_V) randomicamente;

3. Conferimos à variável escolhida uma perturbação que varia de $-\epsilon$ a $+\epsilon$, produzindo um novo estado $s + 1$. Se a variável escolhida é um dos 12 x_j , todo o conjunto é renormalizado à unidade;
4. Perturbações em direção a valores fisicamente impossíveis ($x_i < 0$ ou > 1 , $A_V < 0$) são descartados;
5. Movimentos para regiões com menor χ^2 são **sempre** aceitos;
6. Movimentos para regiões com maior χ^2 são aceitos com probabilidade $e^{-(\chi_{s+1}^2 - \chi_s^2)/2}$, evitando que o algoritmo se prenda em mínimos locais;
7. Todo processo (2–5) é repetido N_{sim} vezes.

No passo número 5 a expressão $e^{-(\chi_{s+1}^2 - \chi_s^2)/2}$ surge, porque deve-se calcular a probabilidade de transição do estado s para o estado $s+1$. Fazemos isso, computando a probabilidade relativa w entre os estados:

$$w = \frac{p(s+1)}{p(s)} = \frac{e^{-\chi_{s+1}^2/2}}{e^{-\chi_s^2/2}} = e^{\frac{-(\chi_{s+1}^2 - \chi_s^2)}{2}} \quad (2.33)$$

Este algoritmo é muito utilizado em Mecânica Estatística (Press *et al.* 1992; MacKay 1999), e produz distribuições cada vez melhores a medida que o número de iterações N_{sim} aumenta. Uma diferença interessante é que em Mecânica estatística é a Energia (e não o χ^2) que é utilizada para aceitar (ou não) as mudanças: se a mudança aumenta a energia do sistema, ela é aceita com uma determinada probabilidade; caso contrário é aceita com probabilidade um. Isto nos dá uma idéia bastante útil sobre a natureza dos erros: **eles desempenham o papel da temperatura dentro deste formalismo**. Quanto maior o erro, maior a “temperatura”, sendo maior a **entropia**. Sendo maior a entropia, aumenta o número de estados acessíveis ao sistema e portanto o número de soluções plausíveis aumenta. Esta analogia nos ajuda a conferir sentido físico à existência dos erros e a entender sua inclusão dentro de um formalismo coerente.

Não existe entretanto uma receita universal para a escolha de valores ótimos para ϵ ou N_{sim} , sendo portanto necessário um pouco de experimentação, que faremos no capítulo seguinte.

Capítulo 3

Testes do algoritmo de Metropolis

No capítulo anterior mostramos o formalismo probabilístico que adotamos e o algoritmo de Metropolis que fará a varredura do espaço de parâmetros. Para escolhermos os melhores valores para o número de iterações N_{sim} e para o parâmetro de visitação ϵ , vamos realizar uma série de testes com um conjunto de galáxias artificiais. Mostraremos também outros resultados que atestam a eficácia do algoritmo em convergir para os valores corretos de \vec{x} e A_V .

O conjunto de galáxias-fictícias é composto por 26 objetos utilizados por Schmitt *et al.* (1991) (tabela 3.1) para testar seu algoritmo de minimização. Como eles não sintetizavam cores, utilizamos um valor $A_V = 2/3$ para a extinção em todas as 26 galáxias.

Esta quantidade aparentemente grande de galáxias-fictícias é necessária porque estamos interessados em resultados estatísticos. Aplicar o formalismo e o algoritmo a poucos objetos não seria útil, podendo inclusive levar a conclusões falsas a respeito do método. É claro que ao longo do capítulo não mostraremos os resultados para todas as galáxias disponíveis. Apenas as utilizamos para “escolher” aquelas que melhor representam ou mostram as propriedades que serão discutidas.

Os métodos de Monte Carlo não nos fornecem uma “receita pronta” para os valores ideais de ϵ e N_{sim} . Em geral, diferentes sistemas apresentam valores ideais diferentes. Por “valor ideal”, nos referimos àqueles que permitem ao algoritmo uma convergência compatível com os limites de tempo de processamento que dispomos e que nossos computadores permitem.

3.1 Relação entre N_{sim} e ϵ no processo Markoviano

O algoritmo precisa amostrar o espaço (\vec{x}, A_V) de maneira a encontrar a região que melhor reproduz os observáveis. Num processo de caminhada aleatória o intervalo de valores que uma dada componente percorre é dado em função de N_{sim} e ϵ por:

$$\Delta x_{Met} \sim \epsilon \sqrt{\frac{N_{sim}}{n_{\star} + 1}} \quad (3.1)$$

Para garantir uma boa amostragem precisamos permitir que os valores dos parâmetros variem entre 0 e 1 algumas vezes. Desta forma, devemos garantir que:

$$\Delta x_{Met} \gtrsim 1 \quad (3.2)$$

Pode-se fazer isso de duas maneiras:

- Poucos passos grandes: N_{sim} pequeno e ϵ grande.
- Muitos passos pequenos: N_{sim} grande e ϵ pequeno.

Conforme discutimos, os resultados são dados em distribuições de probabilidade para as componentes e também para o avermelhamento A_V . Se o algoritmo amostra de uma maneira eficiente o espaço de soluções, esperamos que as curvas de distribuição sejam suficientemente bem concentradas sobre o valor mais provável dos parâmetros. O desvio típico que esperamos para “boas” curvas é da ordem de uns poucos por cento: $\sim 5\text{--}10\%$. Assim, podemos excluir a hipótese de varrer o espaço usando poucas iterações e passos grandes (ϵ grande, $> 10\%$). Como queremos amostrar algo que varia entre $5\text{--}10\%$, devemos adotar a segunda hipótese e varrer o espaço utilizando um bom número de iterações e $\epsilon \sim 1\%$.

Nas figuras desta seção mostramos os testes para as galáxias-fictícias 5, 15 e 20 da tabela 3.1. Consultando a tabela 2.1 vemos que as galáxias 5 e 20 são de população predominantemente velha, enquanto na galáxia 15 predominam populações jovens. Utilizamos 3 valores de N_{sim} : 10^6 , 10^7 e 10^8 (em cores diferentes) e 3 valores de ϵ : $=0.010$, 0.050 e 0.005 (indicados horizontalmente). Para cada galáxia mostramos as curvas de distribuição para 3 das suas componentes mais representativas. Em cada quadro as setas pretas indicam os valores reais (de entrada) para cada componente.

Na figura 3.1 podemos notar que as distribuições deixam de ser suaves a medida que o número de iterações diminui. Além disso, os centros das curvas se distanciam dos valores reais, indicados pelas setas pretas.

Podemos ainda notar que os quadros inferiores (para $\epsilon = 0.005$) são os que apresentam as distribuições mais estreitas e próximas do valor real, para os três valores de N_{sim} . Ou seja: se $\epsilon = 0.005$, os resultados são substancialmente melhores independente do valor de N_{sim} utilizado. O mesmo pode ser verificado para as figuras 3.2 e 3.3. Desta forma vemos que os valores $N_{sim} = 10^8$ e $\epsilon = 0.005$ garantem que o algoritmo visite os estados em passos muito pequenos, aumentando a precisão; e muitas vezes (muitas iterações), aumentando a confiabilidade estatística dos resultados. Computacionalmente estes valores permitem a convergência do método em tipicamente 20 minutos por galáxia num Pentium III 700 MHz.

3.2 Os posteriores projetados

Na figura 3.4 apresentamos as curvas de distribuição para 2 galáxias-fictícias com características distintas: 1 (em rosa e linha contínua) com população predominantemente velha e a 16 (em azul e linha pontilhada), onde dominam as populações jovens. Projetadas em cada quadro estão os posteriores para cada componente e também para o avermelhamento A_V . Os números correspondentes a cada componente estão indicados no canto superior direito e correspondem àqueles da tabela 2.1.

Grupo n° 1: As 26 galáxias-fictícias												
galáxia	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
1	30	20	10	5	20	0	0	10	0	5	0	0
2	0	20	10	5	30	0	0	20	0	10	0	5
3	0	20	10	5	0	30	0	0	20	0	10	5
4	0	0	10	5	0	30	20	0	20	0	10	5
5	50	20	5	0	20	0	0	5	0	0	0	0
6	0	50	20	5	0	20	0	0	5	0	0	0
7	0	20	5	0	50	0	0	20	0	5	0	0
8	0	20	5	0	0	50	0	0	20	0	5	0
9	70	10	5	0	10	0	0	5	0	0	0	0
10	0	70	10	5	0	10	0	0	5	0	0	0
11	0	10	5	0	70	0	0	10	0	5	0	0
12	0	10	5	0	0	70	0	0	10	0	5	0
13	10	10	10	0	40	0	0	20	0	10	0	0
14	0	10	10	10	0	40	0	0	20	0	10	0
15	0	10	10	0	0	10	0	0	40	0	20	10
16	0	10	10	0	0	10	0	0	20	0	40	10
17	0	10	40	40	0	10	0	0	0	0	0	0
18	0	30	20	10	0	10	5	0	10	0	10	5
19	0	10	20	50	0	0	0	0	0	0	20	0
20	0	10	20	50	0	0	0	0	10	0	10	0
21	0	10	10	10	0	5	5	0	20	5	30	5
22	0	0	10	10	0	0	10	0	30	10	20	10
23	0	0	0	0	20	0	0	50	0	20	0	10
24	0	0	0	0	0	10	0	0	30	0	50	10
25	0	0	0	0	0	0	0	20	0	70	0	10
26	0	0	0	0	0	0	0	0	70	0	20	10

Tab. 3.1: As 26 galáxias-fictícias apresentadas em Schmidt *et al.* (1991). Horizontalmente podem ser lidos o número das galáxias e as frações do fluxo total (normalizadas em $\lambda 5870$ Å, representadas em porcentagem) das suas respectivas componentes. Os números das componentes correspondem àqueles indicados na tabela 2.1.

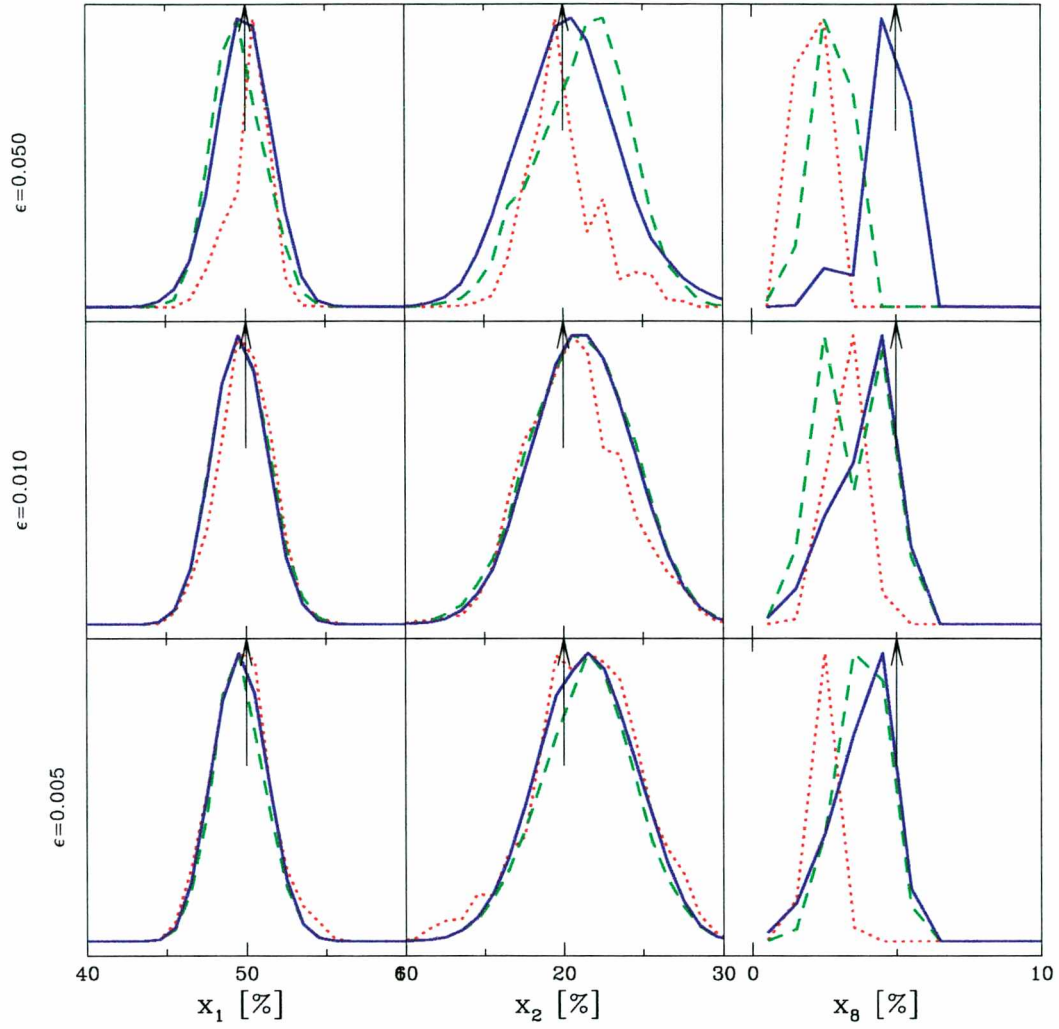


Fig. 3.1: Distribuições de probabilidade para as componentes #1, #2 e #8 da galáxia 5, que respondem por 75% de sua luz. Estão representados 3 valores de N_{sim} : curva em vermelho (pontilhada)= 10^6 , verde (tracejada)= 10^7 e azul (contínua)= 10^8 . Horizontalmente podem ser lidos os valores de ϵ correspondentes.

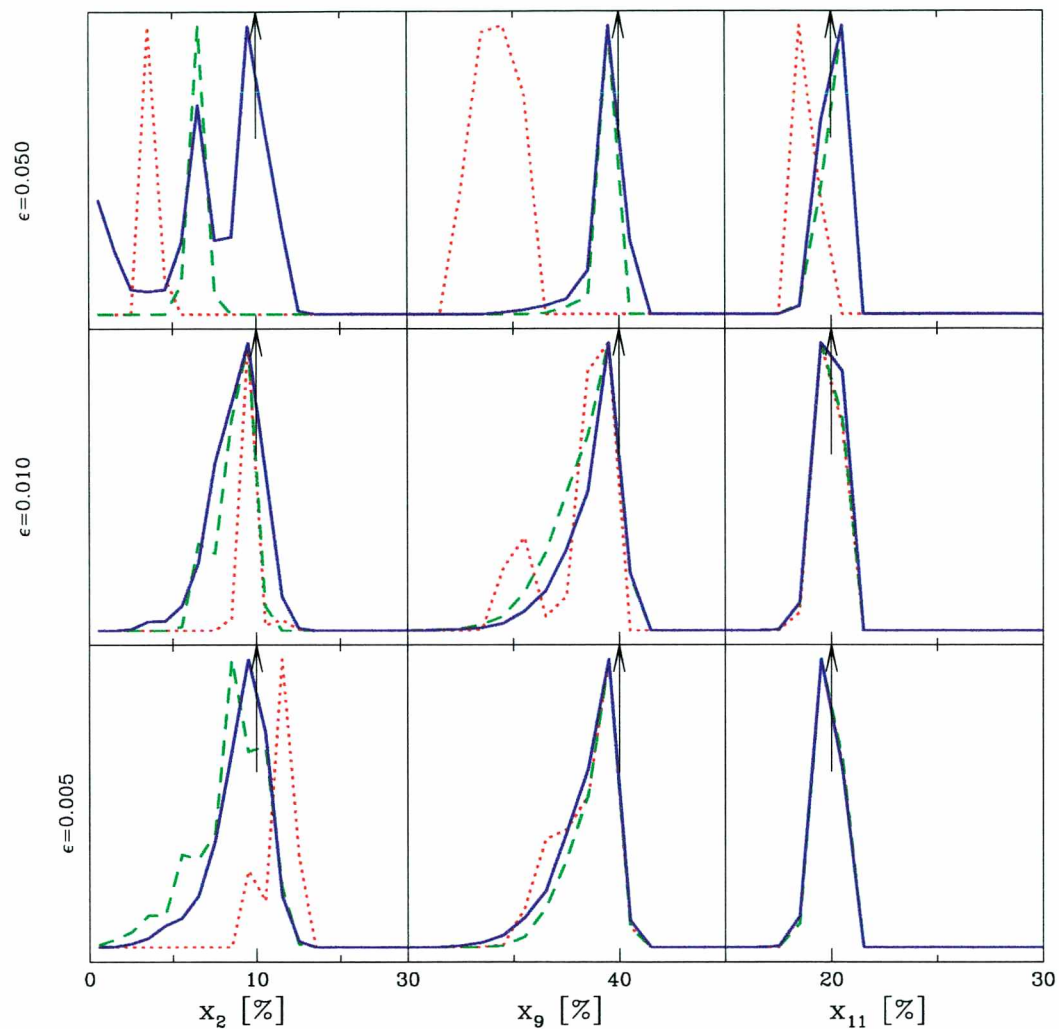


Fig. 3.2: Distribuições de probabilidade para a galáxia 15. Estão representadas as componentes #2, #9 e #11, que respondem por 70% da luz deste objeto. As convenções de cores são as mesmas da figura anterior.

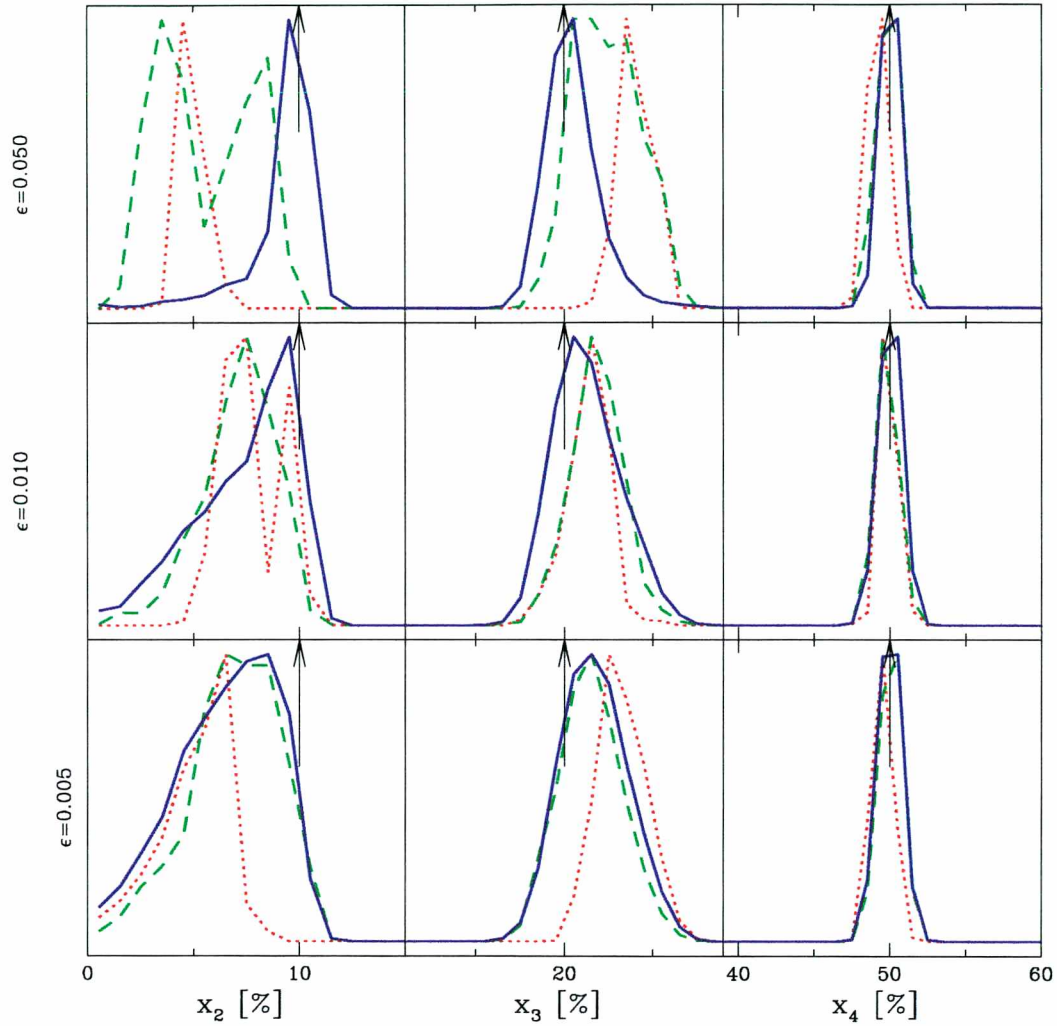


Fig. 3.3: Curvas de distribuição de probabilidade para a galáxia 20. Estão representadas as componentes #2, #3 e #4, que respondem por 80% da luz deste objeto. As convenções de cores são as mesmas das 2 figuras anteriores.

As distribuições foram obtidas começando-se pelo ponto $x_1 = x_2 = \dots = x_{12} = 1/12$ e $A_V = 0.5$. Utilizou-se o conjunto A de observáveis ¹ composto por 9 linhas e 7 cores e a qualidade do espectro $(S/N)_{5870} = 1000$. É claro que se trata de um caso bastante idealizado, mas serve de teste para o algoritmo no limite onde dispõe-se de todos os observáveis da base e também de um espectro de alta qualidade. A figura ilustra também a forma dos resultados que serão apresentados no capítulo seguinte. Podemos notar claramente que o algoritmo converge para a região mais provável do espaço de soluções. Os valores de entrada (reais) estão indicados por setas. Podemos notar que as componentes mais velhas (x_1 , x_2 e x_3) da galáxia 1 possuem posteriores com maior desvio. Este efeito será discutido no capítulo 4.

3.3 A evolução do algoritmo de Metropolis

Para ilustrar a evolução do algoritmo de Metropolis ao longo do espaço de parâmetros, mostramos na figura 3.5 como o algoritmo varre o espaço de parâmetros na tentativa de encontrar os melhores valores para eles. Para simplificar a visualização, estão representadas apenas 6 componentes e um para cada 10^4 passos é mostrado. Nos quadros superiores estão mostrados o χ^2 e dois observáveis: a largura equivalente W_{CaIIK} e a cor C_{4020}/C_{5870} . A qualidade dos dados nesta simulação é $S/N = 60$ e foi utilizado o conjunto A composto por 16 observáveis. Os valores de entrada estão indicados em vermelho no canto direito em cada quadro.

Podemos ver claramente como o método visita estados bastante afastados dos valores de entrada. São notáveis também os efeitos de compensação entre componentes de mesma idade ao longo da “caminhada” pelo espaço de soluções: para cada pico no valor de x_2 observamos vales no valor de x_1 , o mesmo ocorrendo para x_8 e x_9 . Este efeito será discutido em detalhes no capítulo 4. Ao contrário do que ocorre com os parâmetros, os observáveis sintéticos oscilam sempre em torno dos valores de entrada, indicando que o algoritmo é mais preciso para determinar os observáveis do que os parâmetros.

¹ Os grupos de observáveis serão apresentados no capítulo 4.

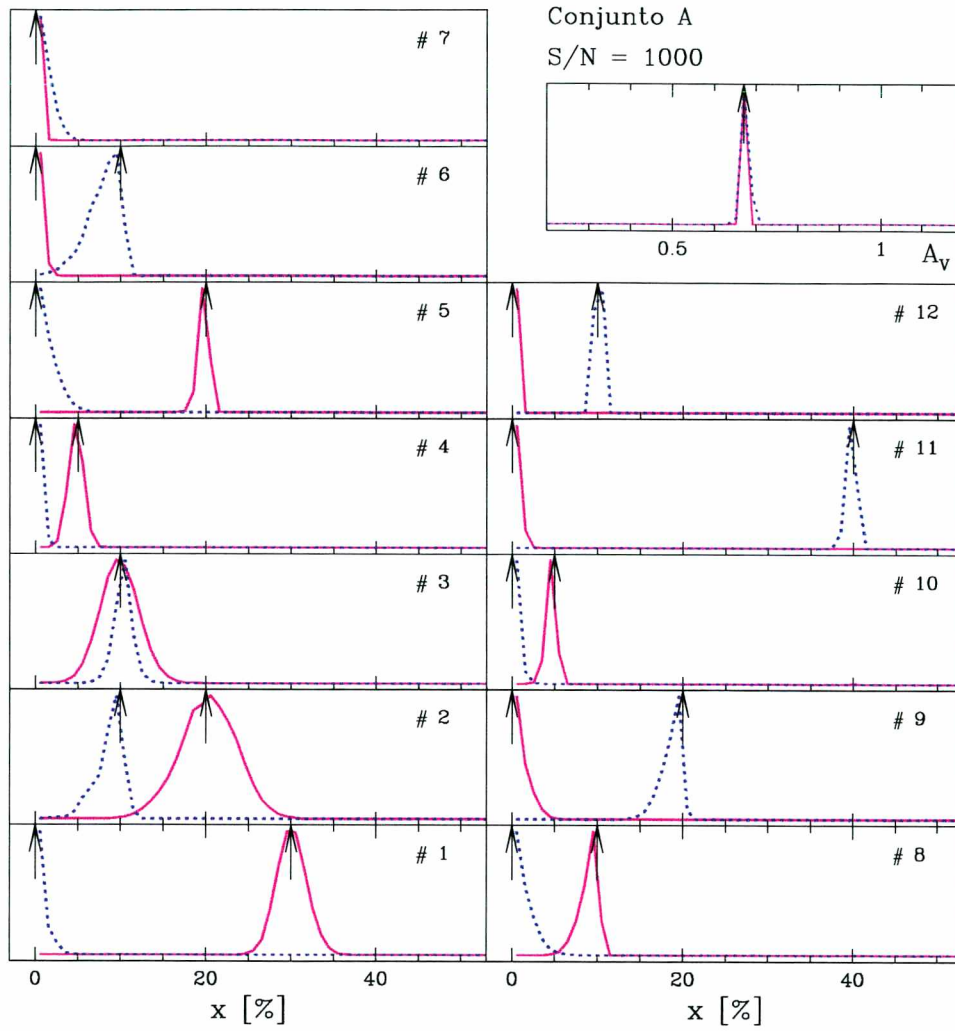


Fig. 3.4: Distribuição de probabilidades para todas as componentes e o avermelhamento para as galáxias 1 (em rosa, contínua) e 16 (azul, pontilhada). A qualidade do espectro e o conjunto de observáveis utilizados estão indicados

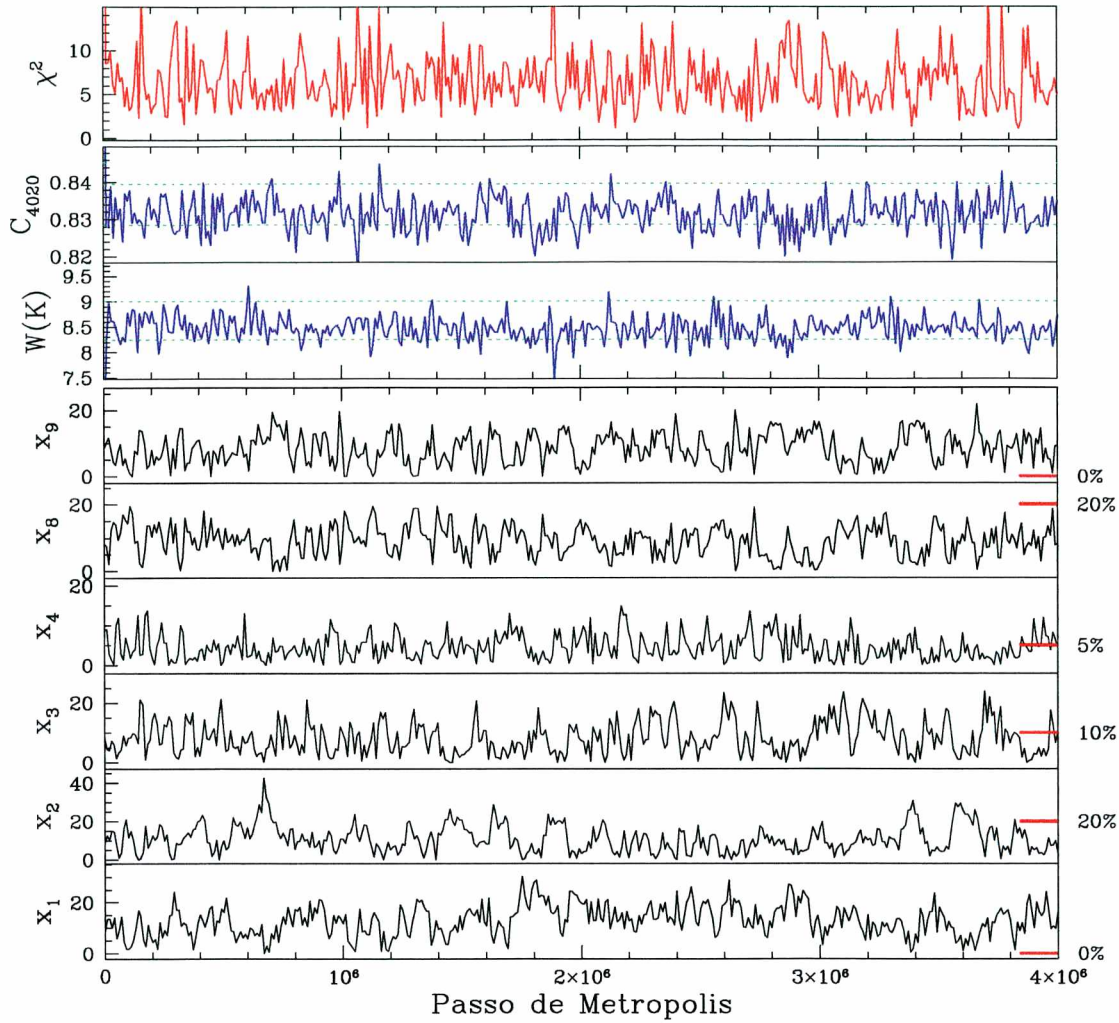


Fig. 3.5: Ilustração da caminhada do algoritmo de Metropolis pelo espaço de soluções para a galáxia-fictícia 2. Os valores de entrada estão indicados por traços vermelhos na extremidade direita de cada painel. O primeiro painel indica a evolução do χ^2 para cada estado visitado, os dois painéis seguintes referem-se aos observáveis indicados e os seis painéis inferiores indicam a evolução dos x_i . As linhas pontilhadas no 2º e 3º painéis indicam os desvios $\sigma = \pm 1$ em torno do valor de entrada. Foi utilizado o conjunto A (todos os observáveis da base) e $S/N = 60$.

3.4 O limite da “temperatura infinita”

Vamos agora imaginar uma situação totalmente oposta daquela da figura 3.4: dados de pouca qualidade (espectro ruidoso). Neste limite, dos erros tendendo ao infinito ($S/N \rightarrow 0$), a verossimilhança perde seu poder de resolução e fica constante ao longo de todo o espaço; ou seja, todos os estados são equiprováveis. Este regime é conhecido em Mecânica Estatística como “temperatura infinita” e concorda com a analogia que fizemos entre a Termodinâmica e a natureza dos erros na seção 2.5.2. A equação para a verossimilhança pode ser resolvida analiticamente:

$$P(x_i|D, \sigma, I) \sim \int \dots \int \text{const} \, dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_{11} dA_V \quad (3.3)$$

O termo da direita na equação acima é proporcional ao volume no espaço de parâmetros no qual a componente x_i se encontra entre x_i e $x_i + dx_i$. Para a componente 1 este volume pode ser escrito como:

$$v(x_1) \equiv \int \dots \int dx_2 \dots dx_{11} = \int_0^{1-x_1} dx_2 \int_0^{1-x_1-x_2} dx_3 \dots \int_0^{1-\sum_{i=1}^{10} x_i} dx_{11} \quad (3.4)$$

A última destas integrais fica:

$$1 - \sum_{j=1}^{10} x_j = \left(1 - \sum_{j=1}^9 x_j \right) - x_{10} \equiv a_{10} - x_{10} \quad (3.5)$$

Substituindo na integral para x_{10} obtem-se:

$$\int_0^{1-\sum_{i=1}^9 x_i = a_{10}} (a_{10} - x_{10}) dx_{10} = a_{10} \int dx_{10} - \int x_{10} dx_{10} = (a_{10}x_{10} - x_{10}^2/2) \quad (3.6)$$

aplicando os limites, $0 - a_{10}$, temos:

$$(a_{10}a_{10} - a_{10}^2/2) = a_{10}^2/2 \quad (3.7)$$

que pode ser escrito como $(a_9 - x_9)^2/2$, onde $a_9 = a_8 - x_8$. A regra geral para os integrandos subsequentes é então:

$$a_n = a_{n-1} - x_{n-1} \quad (3.8)$$

A integral em dx_9 fica $a_9^3/6$. Se repetirmos esta operação para a integral em dx_8 , obtemos $a_8^4/24$. Assim, vemos que $v(x_1)$ é dado pela equação:

$$v(x_i) = \frac{a_2^{10}}{10!} = \frac{(1 - x_1)^{10}}{10!} \quad (3.9)$$

Portanto a verossimilhança é proporcional a $(1 - x_1)^{10}$. Desta distribuição obtem-se que o valor médio é $\bar{x}_i = 1/12$, tal como intuitivamente se espera. Além disto, como existem muitos estados na região $x_i \sim 0$, os posteriores tem seu pico em $x_i = 0$.

Na Figura 3.6 são mostrados os posteriores para as componente 2, 6 e 9 da galáxia-fictícia 4 da tabela 3.1, calculados para uma razão S/N de 1, ou seja, para um espectro extremamente ruidoso. Vê-se que **todos** os posteriores se sobrepõem, apesar das componentes 2, 6 e 9 terem valores de entrada bem diferentes: $x_2 = 0\%$, $x_6 = 30\%$ e $x_9 = 20\%$. Note também que nosso método de amostragem produz uma distribuição de probabilidades que se aproxima satisfatoriamente daquela prevista (linha sólida). Os valores médios obtidos numericamente são $\overline{x}_2 = 8.40$, $\overline{x}_6 = 8.40$ e $\overline{x}_9 = 8.30\%$, em excelente acordo com a previsão $\overline{x}_i = 8.33\%$.

Desta forma, comprovamos a eficiência do algoritmo de Metropolis em amostrar apenas as regiões relevantes do espaço de parâmetros. No capítulo seguinte, os resultados da síntese são mostrados e outros efeitos são discutidos.

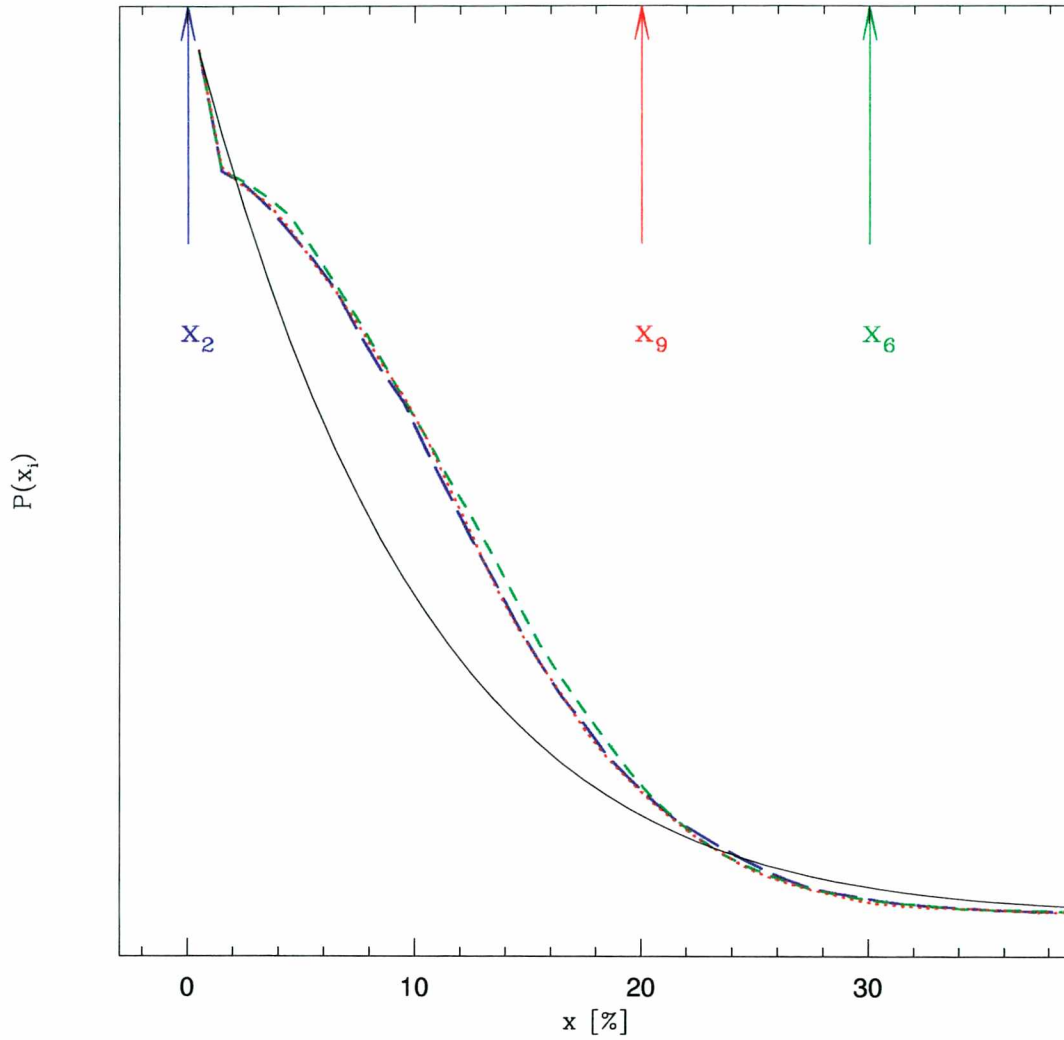


Fig. 3.6: Curvas de distribuição para as componentes 2 (azul), 6 (verde) e 9 (vermelho) da galáxia 4. Notar como os posteriores tendem para o valor teórico $\bar{x}_i = 8.33\%$, indicado pela curva em preto, no limite da temperatura infinita. As setas indicam os valores de entrada das componentes.

Capítulo 4

Testes da base com o algoritmo de Metropolis

No capítulo anterior testamos o **Algoritmo de Metropolis** e decidimos que os valores: $\epsilon = 0.005$ e $N_{sim} = 10^8$, são os que produzem as melhores distribuições, possibilitando a convergência do método.

Neste capítulo vamos mostrar os resultados da síntese para as mesmas galáxias-fictícias da tabela 3.1. Além destas, em algumas figuras mostramos resultados para outras 100 galáxias-fictícias, construídas a partir da base, onde sempre domina uma das componentes. Utilizamos espectros com qualidades variadas e utilizamos 3 diferentes conjuntos de observáveis:

- Conjunto A: Todos os 16 observáveis da base: 9 larguras equivalentes e 7 cores
- Conjunto B: Apenas as 9 larguras equivalentes
- Conjunto C: 4 W_i 's (CaII K, CN, banda-G, e MgI+MgH) e 4 cores (3660, 4020, 4510 e 6630 Å)

Desta forma, os resultados são apresentados em função dos **erros observacionais** e do **grupo de observáveis** utilizado. Isto vai nos permitir testar a base, entender se a inclusão das cores melhora os resultados e determinar quantitativamente a influência da relação Sinal-Ruído (S/N) nas curvas de distribuição.

4.1 O efeito da qualidade do espectro: A relação S/N

É bastante intuitivo que os erros observacionais, que refletem a qualidade do espectro, influenciam os resultados da síntese. Isto porque o espectro é a única fonte de informações que utilizamos neste método. Qualitativamente esperamos que espectros ruins não permitam bons resultados, ao passo que espectros de melhor qualidade possam ser mais úteis. Mas como quantificar o efeito dos erros?

Nesta seção vamos mostrar efetivamente como os erros afetam as distribuições de probabilidade.

A figura 4.1 mostra as distribuições de probabilidade para três relações sinal/ruído ($(S/N)_{5870}$) diferentes. Sendo: linha vermelha (pontilhada)=10, verde(tracejada)=100

e azul (contínua)=1000. Esta figura mostra os resultados para a galáxia 3, cujo vetor de populações pode ser visto na tabela 3.1. Nesta galáxia dominam populações de idade intermediária (10^8 – 10^9 anos).

Podemos notar que para erros maiores (menor S/N) as distribuições se tornam mais largas, como intuitivamente se espera. Além disso, a medida que os erros se tornam maiores, a média das distribuições tendem a se deslocar para a esquerda ou para a direita. Notamos entretanto que componentes de mesma idade se compensam: 1 e 2 (10^{10} anos); 6 e 7 (10^9 anos); 8 e 9 (10^8 anos). Este efeito é o mesmo apontado na seção 3.3 e que pode ser visualizado na figura 3.5.

O efeito dos erros pode também ser visto nas figuras 4.2 para a galáxia 21, com população bastante mesclada, mas onde dominam populações jovens; e 4.3, referente à galáxia 6 onde as populações velhas são mais fortes.

Na figura 4.4 mostramos uma representação gráfica da “matriz de correlação” x_i X x_j , onde os pontos azuis representam pontos visitados pelo algoritmo de Metropolis. Podemos notar os efeitos de correlação para as componentes da mesma galáxia da figura 4.1. Podemos ver “nuvens” que denotam as anti-correlações entre elementos de mesma idade. O espalhamento aumenta com a diminuição da qualidade do espectro. Neste exemplo, $S/N = 300$.

A explicação deste **efeito de compensação** pode ser obtida analisando a base: seus elementos deveriam ser a princípio linearmente independentes mas o que se nota é que alguns dos elementos podem ser “construídos” utilizando-se outros elementos da própria base. Com o aumento dos erros, as diferenças entre o elemento original e o construído são insignificantes, dando origem a um efeito que poderíamos chamar de **dependência linear estatística** entre as componentes. Isto faz com que no limite da temperatura infinita, discutido na seção 3.4, algumas componentes (principalmente as de mesma idade) se tornem praticamente indistinguíveis. Isto explica o porquê das redistribuições presentes nas figuras 4.1, 4.2 e 4.3 a medida que cai a qualidade do espectro.

A figura 4.5 ilustra este fato, onde mostramos o resultado da síntese dos próprios elementos da base. Estão representados os resultados para 3 diferentes relações S/N (100 (branco), 30 (preto) e 10 (verde)) e utilizando o conjunto A de observáveis. Cada painel(B1,B2,...B12) corresponde a uma galáxia-fictícia cuja luz provém de um único elemento da base. Podemos notar que mesmo para um espectro de ótima qualidade como $S/N = 100$ algumas componentes são bem sintetizadas com contribuições de outras componentes. Em especial, x_2 e x_3 que podem ser sintetizadas com combinações das componentes x_1 e x_4 e x_6 que pode ser obtida por combinações de x_5 , x_7 .

Na figura 4.6 plotamos os valores de entrada para as componentes contra seus valores médios, para 100 galáxias-fictícias. Os círculos azuis indicam $S/N = 1000$ e as cruzeiras vermelhas $S/N = 60$. Um conjunto diferente de 100 galáxias-fictícias foi utilizado. Podemos ver que mesmo para dados muito bons (indicados pelas cruzeiras), é muito difícil recuperar as populações estelares. Apenas para espectros não realísticos (círculos) os parâmetros podem ser recuperados com exatidão, quebrando a dependência linear estatística. Entretanto, o algoritmo sintetiza os observáveis e estes podem ser recuperados com extrema precisão.

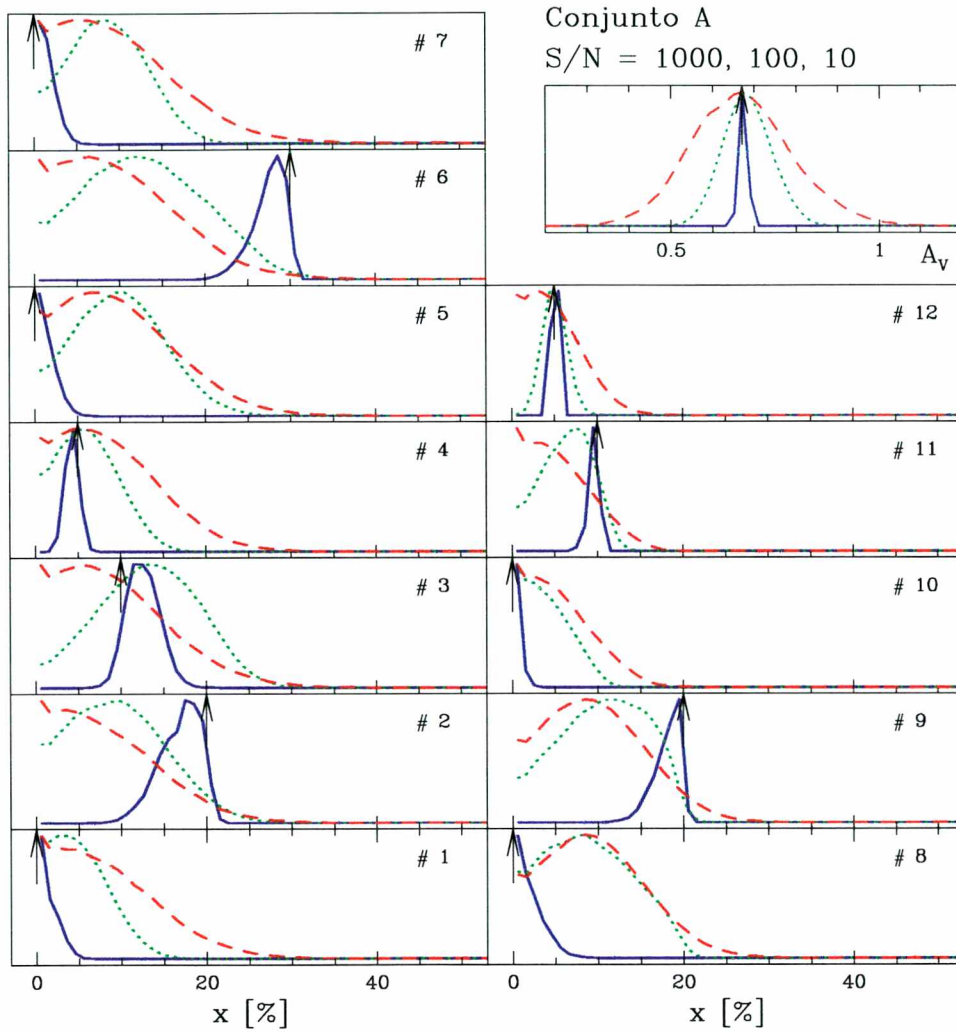


Fig. 4.1: Curvas de distribuição para a galáxia-fictícia 3 (com população intermediária dominante) mostrando o efeito dos erros observacionais nos posteriores do vetor \vec{x} e A_V . O conjunto de observáveis está indicado e utilizamos 3 valores de S/N : 1000 (linha azul contínua), 100 (linha verde tracejada) e 10 (linha vermelha pontilhada). Vemos que os posteriores das componentes mais importantes se redistribuem entre outras componentes a medida que o ruído aumenta.

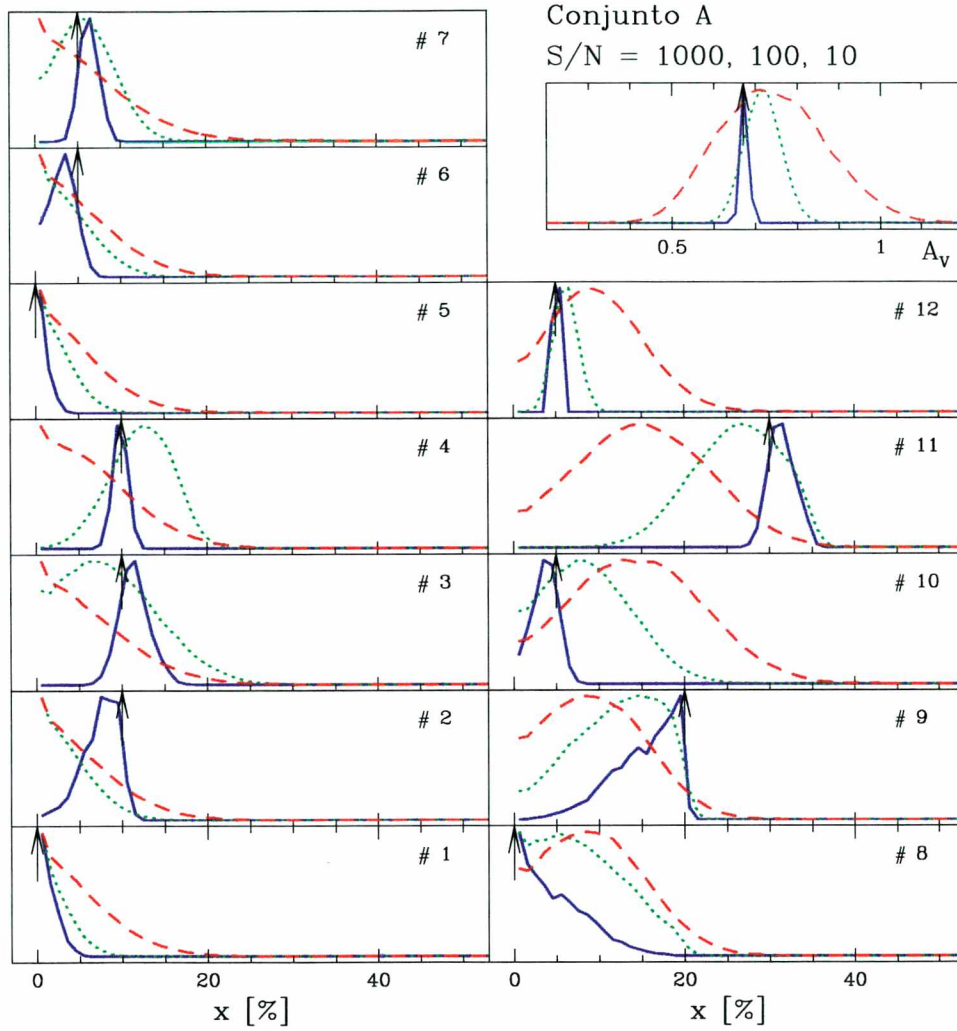


Fig. 4.2: Distribuições de probabilidade para a galáxia-fictícia 21, onde dominam populações jovem, mostrando o efeito dos erros no vetor \vec{x} e A_V . O conjunto de observáveis está indicado e utilizamos 3 valores de S/N , seguindo a mesma convenção de cores da figura acima. O mesmo efeito de distribuição da figura anterior está presente.

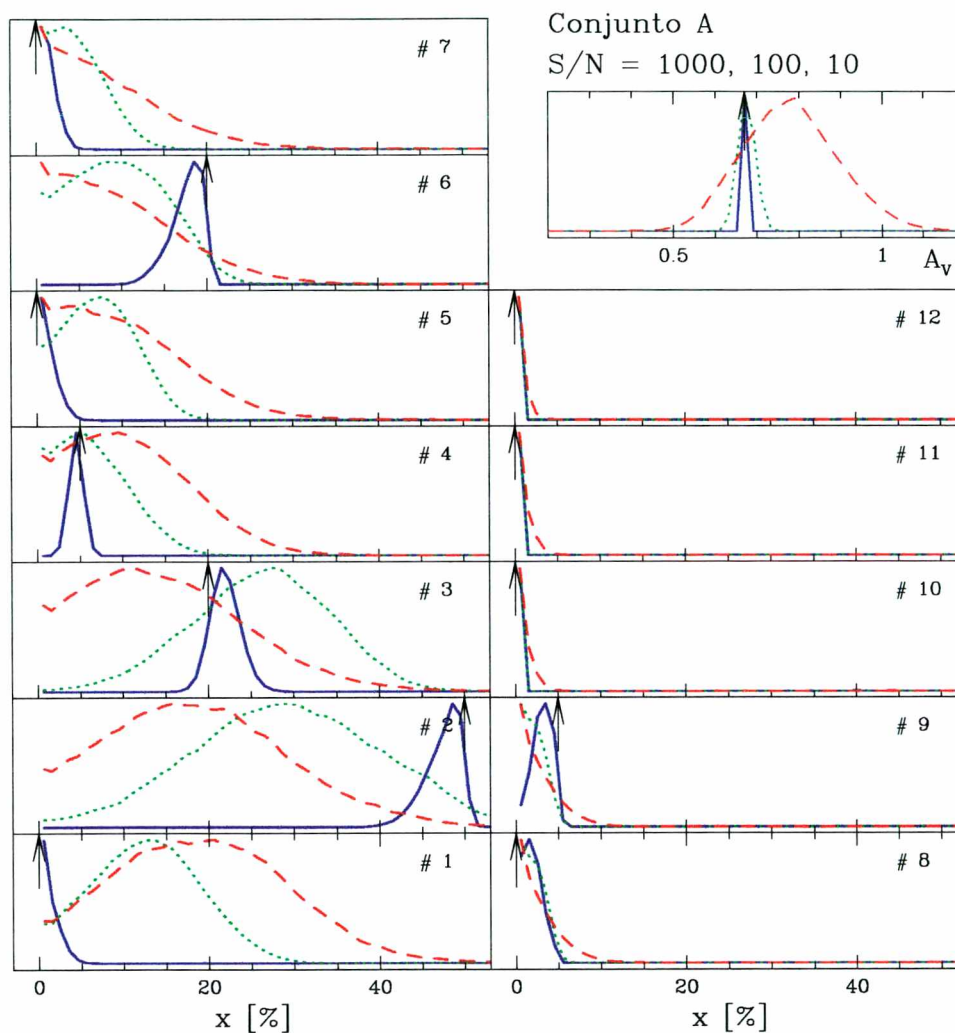


Fig. 4.3: O mesmo da figura anterior, inclusive as convenções de cores, porém para a galáxia-fictícia 6 (relativamente velha). Notar a redistribuição das componentes mais velhas a medida que cai a qualidade do espectro

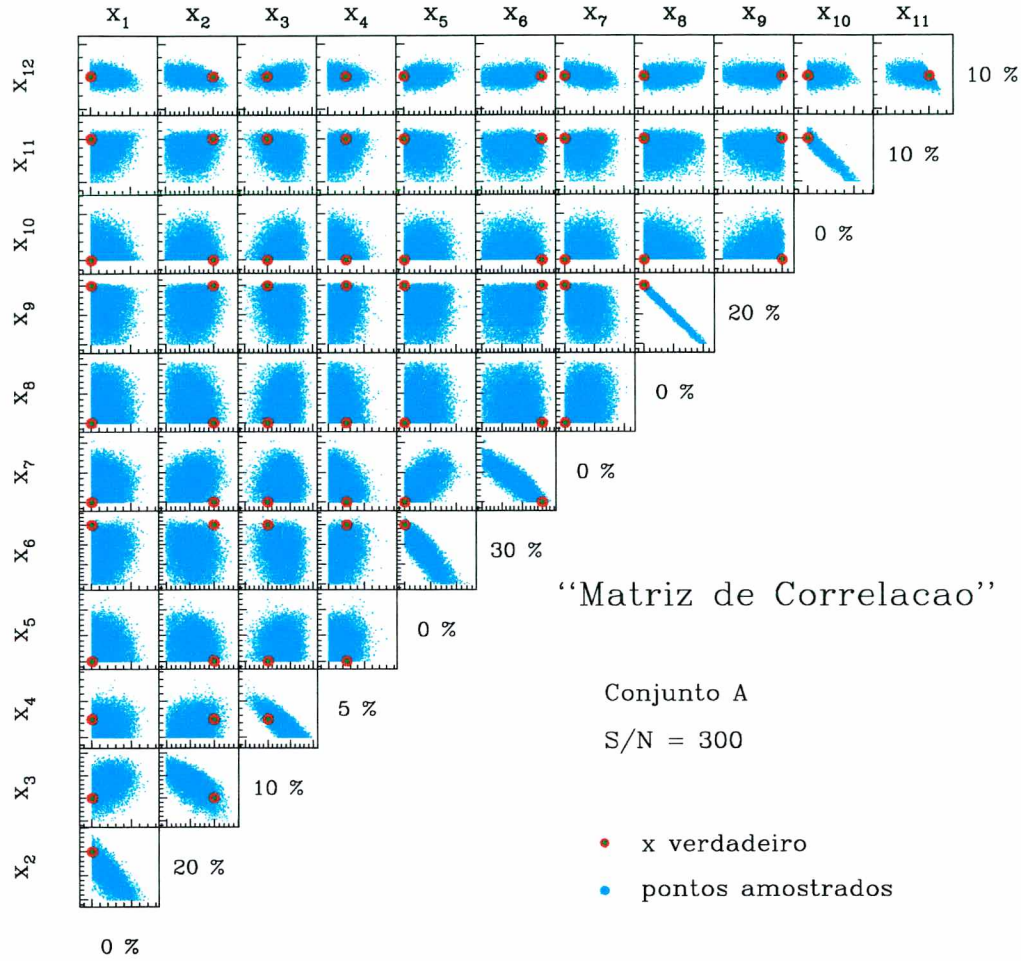


Fig. 4.4: Matriz de correlação para as componentes da galáxia-fictícia 3 (figura 4.1). A qualidade do espectro e o conjunto de observáveis estão indicados. Em cada painel $x_i \times x_j$ estão representados apenas 1 em cada 10^4 passos de uma simulação com $N_{sim} = 10^7$. Os valores reais estão indicados na parte inferior de cada coluna e também pelos círculos vermelhos. Podemos notar a forte anti-correlação entre elementos de mesma idade: x_1 e x_4 , x_5 e x_7 , x_{10} e x_{11} .

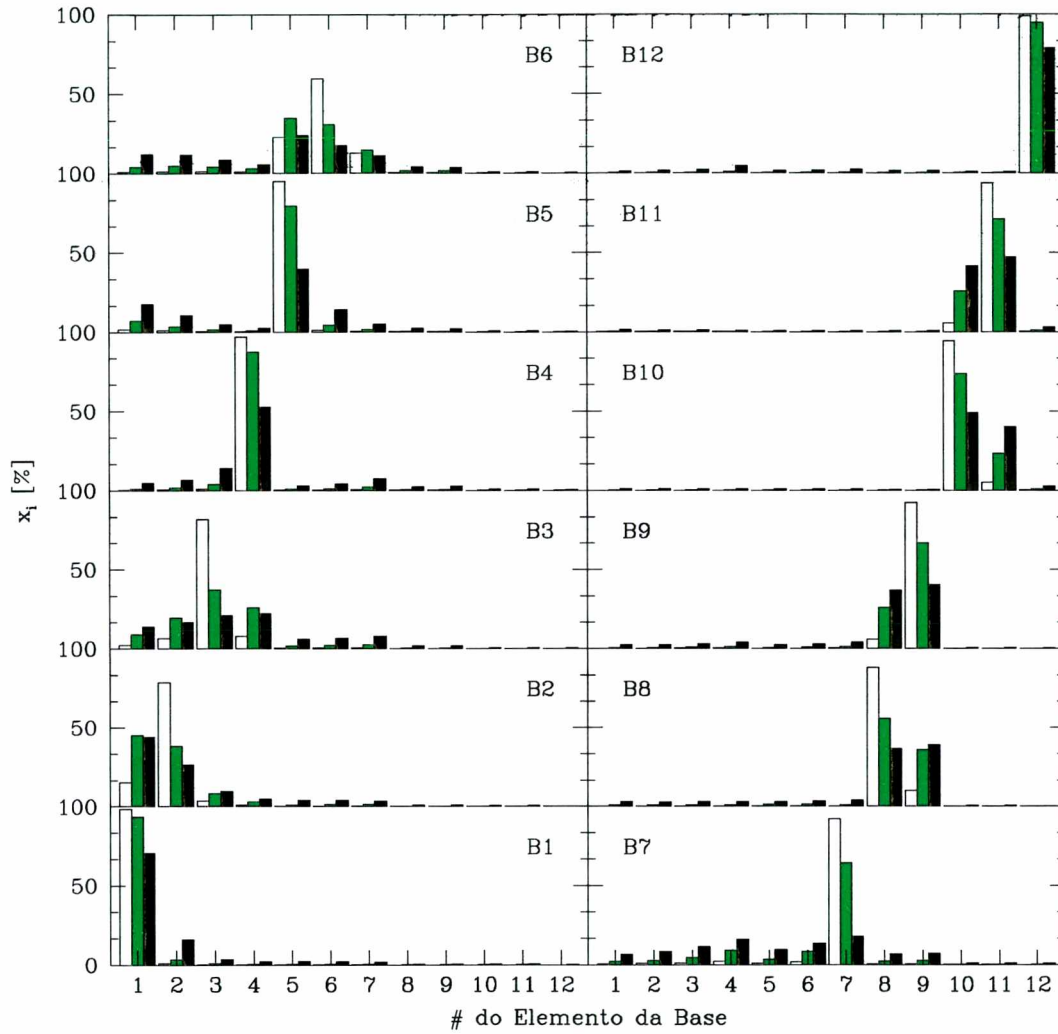


Fig. 4.5: Síntese de 12 galáxias fictícias (B1, B2,...,B12) onde apenas uma das componentes responde por 100% da sua luz. Isto equivale a sintetizar os elementos da base. No eixo vertical estão indicados os valores de x_{medio} e no outro eixo estão indicados os elementos da base. Três relações sinal/ruído estão representadas: 100 (branco), 30 (preto) e 10 (verde). Idealmente esperamos que os histogramas estivessem concentrados no observável que o gerou. No entanto vemos que alguns elementos são sintetizados por combinações de outros elementos da base. Isto ocorre principalmente com as componentes 2, 3 e 6.

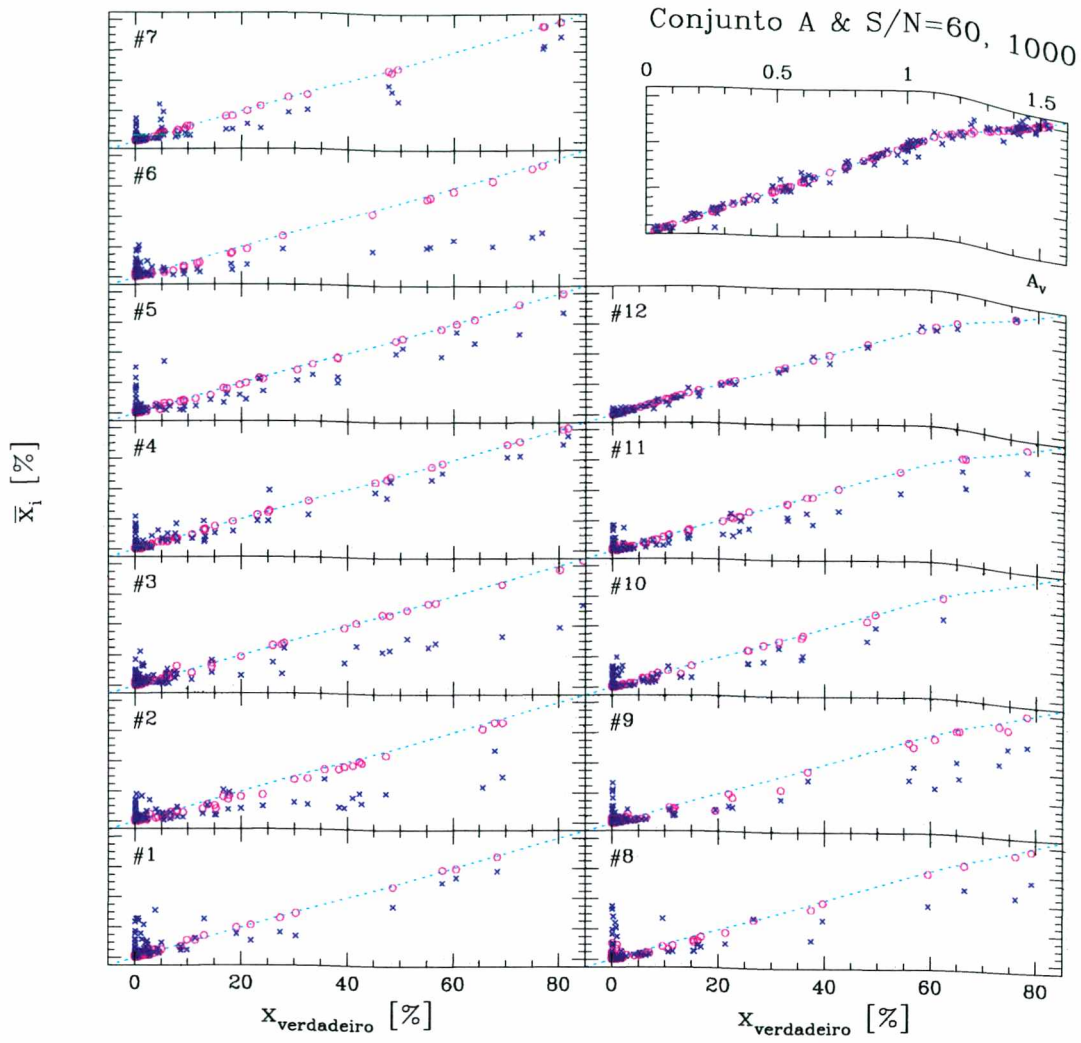


Fig. 4.6: Valores de entrada contra valores médios, para 100 galáxias-fictícias. Estão presentes duas relações S/N : círculos rosas indicam $S/N = 1000$ e as cruzes azuis $S/N = 60$. Notamos que mesmo para espectros bons ($S/N \sim 60$), é difícil recuperar as populações estelares.

Mostramos na figura 4.7 um gráfico dos observáveis sintéticos contra os observados para várias relações S/N . O melhor modelo encontrado durante o processo de síntese está também representado, onde podemos notar uma concordância ainda melhor. Vemos que os observáveis são precisamente recuperados mesmo para espectros não idealizados.

Ressaltamos que estas limitações não são devido ao método utilizado, mas sim à estrutura interna da base. Podemos desta forma confiar na capacidade do algoritmo em recuperar os observáveis, mas não nos detalhes do vetor de populações. Nota-se que a dependência linear estatística ocorre principalmente entre elementos **de mesma idade**, sugerindo que a distribuição de idades pode ser recuperada se os resultados forem agrupados em idade; apesar da existência de redistribuição entre componentes mais velhos (10^9 e 10^{10} anos) devido à degenerescência entre idade e metalicidade. (Ver seção 4.4).

4.2 O efeito dos conjuntos de observáveis

Vamos agora mostrar como o número de observáveis sendo sintetizado se reflete na precisão do método.

Qualitativamente esperamos que quanto maior for o número de informações que dispomos, melhores serão os resultados. Temos 12 parâmetros (considerando o vínculo de normalização) a serem determinados (“variáveis”) e os observáveis, que funcionam como “equações”, e reduzem a degenerescência algébrica.

Nas figuras 4.8 a 4.10 mostramos os resultados para as galáxias 3, 21 e 6 da tabela 3.1 (para efeito de comparação as mesmas apresentadas na seção anterior). A relação S/N está indicada e os resultados estão superpostos para os três grupos de observáveis: conjunto A em azul (linha contínua), conjunto B em verde (tracejada) e conjunto C em vermelho (pontilhada). As setas pretas indicam os valores reais de cada componente e para o avermelhamento.

Podemos notar que para os grupos com menos informações há um maior desvio das distribuições. Comparando as curvas produzidas pelos conjuntos A e B, vemos que apesar de se incluir um parâmetro extra com a introdução das cores, os resultados são muito melhores quando são sintetizadas em conjunto com as larguras equivalentes. Felizmente, ao contrário dos x_i , a extinção pode sempre ser recuperada com precisão; além de melhorar os resultados para o vetor \vec{x} .

Comparando as curvas A, B e C vemos que o aumento do número de observáveis na síntese contribui para aumentar os vínculos do problema e consequentemente a precisão dos resultados é melhorada. Desta forma, a redução dos observáveis produz o mesmo resultado da redução da qualidade do espectro.

4.3 Resultados agrupados em idade

Ficou claro a partir das duas 2 seções anteriores que tanto a redução da qualidade do espectro quanto a utilização de conjuntos de observáveis com menos informações tendem a espalhar as curvas de distribuição entre componentes de mesma idade. Isto sugere que o agrupamento dos resultados por idade possa nos fornecer melhores

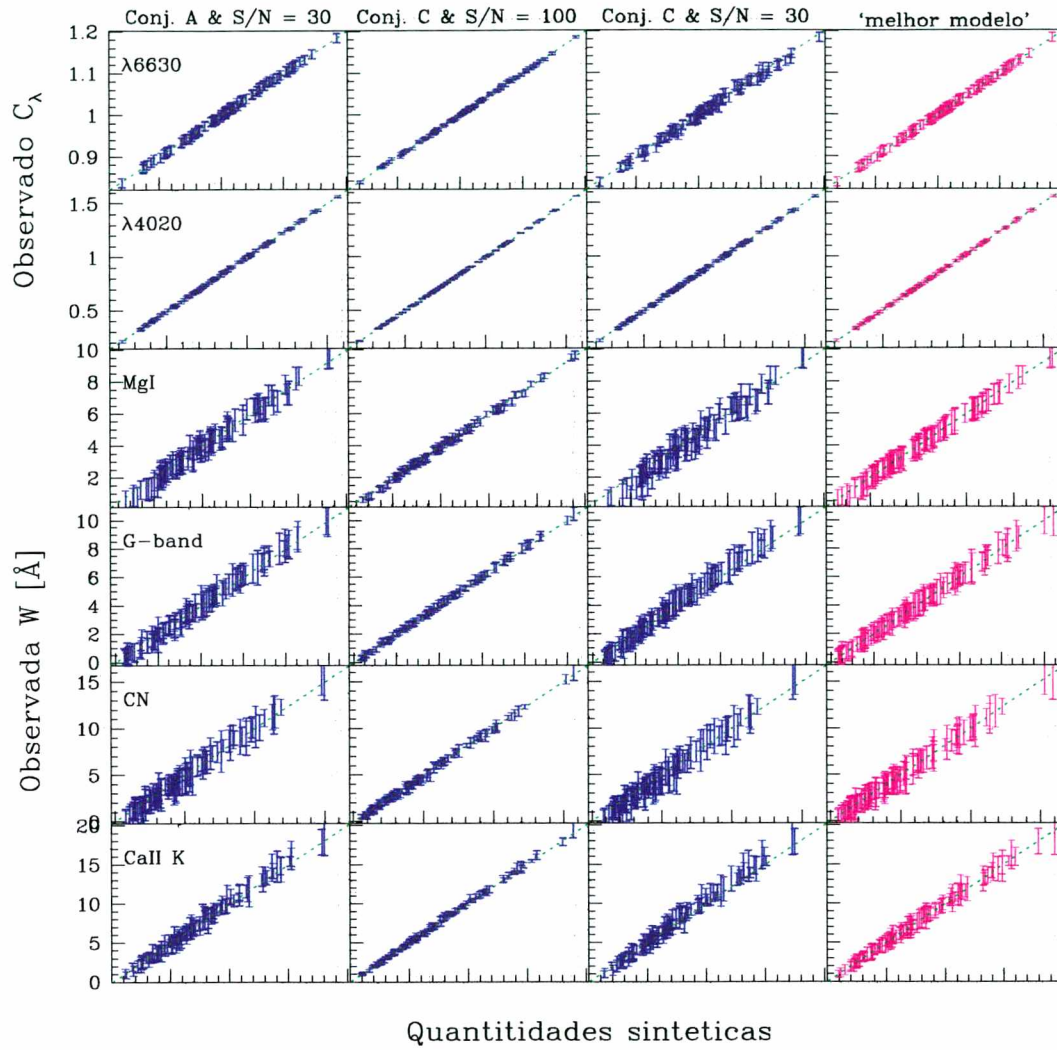


Fig. 4.7: Valores de entrada contra sintéticos (x_{medio}) para as 100 galáxias-fictícias e as combinações de conjunto de observáveis e S/N indicadas. Apenas alguns observáveis estão mostrados, mas a concordância é igualmente boa para todos os outros. Na última coluna esta representado o melhor modelo encontrado durante a simulação.

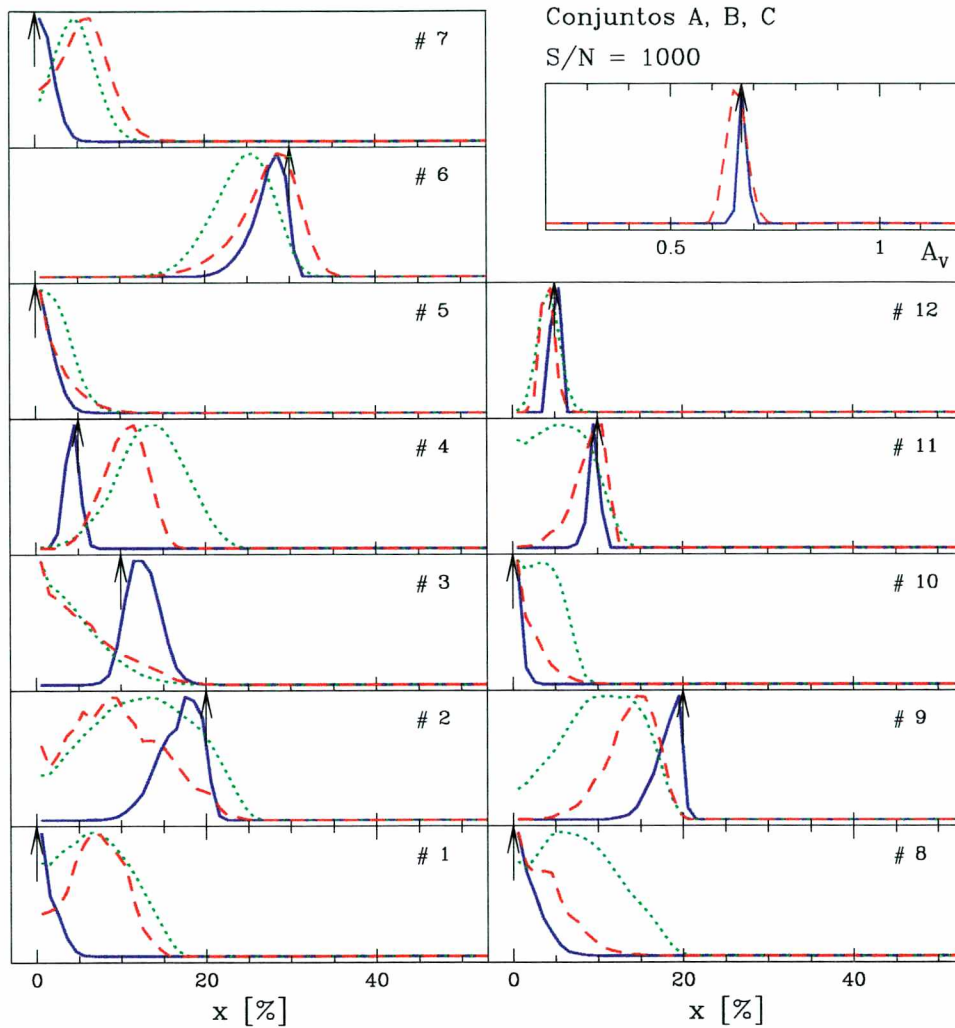


Fig. 4.8: Resultados para a galáxia-fictícia 3 (com população intermediária dominante) mostrando o efeito dos conjuntos de observáveis nos posteriores do vetor \vec{x} e A_V . O valor de S/N está indicado e utilizamos 3 conjuntos de observáveis: A (linha azul contínua), B (linha verde tracejada) e C (linha vermelha pontilhada). Podemos notar como as distribuições ficam menos concentradas para grupos com menos observáveis.

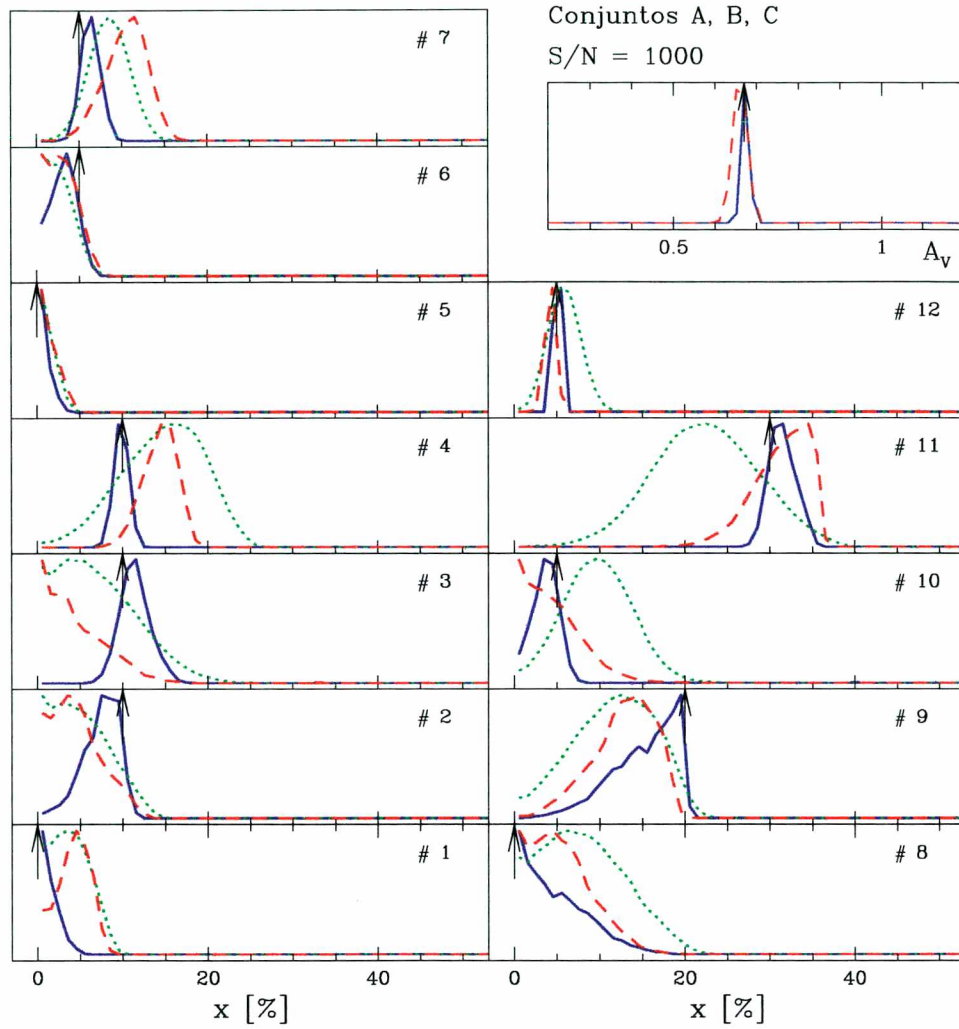


Fig. 4.9: Resultados para a galáxia-fictícia 21 (com população relativamente jovem) mostrando o efeito dos conjuntos de observáveis nos posteriores do vetor \vec{x} e A_V . O valor de S/N está indicado e utilizamos 3 conjuntos de observáveis, seguindo a mesma convenção da figura anterior.

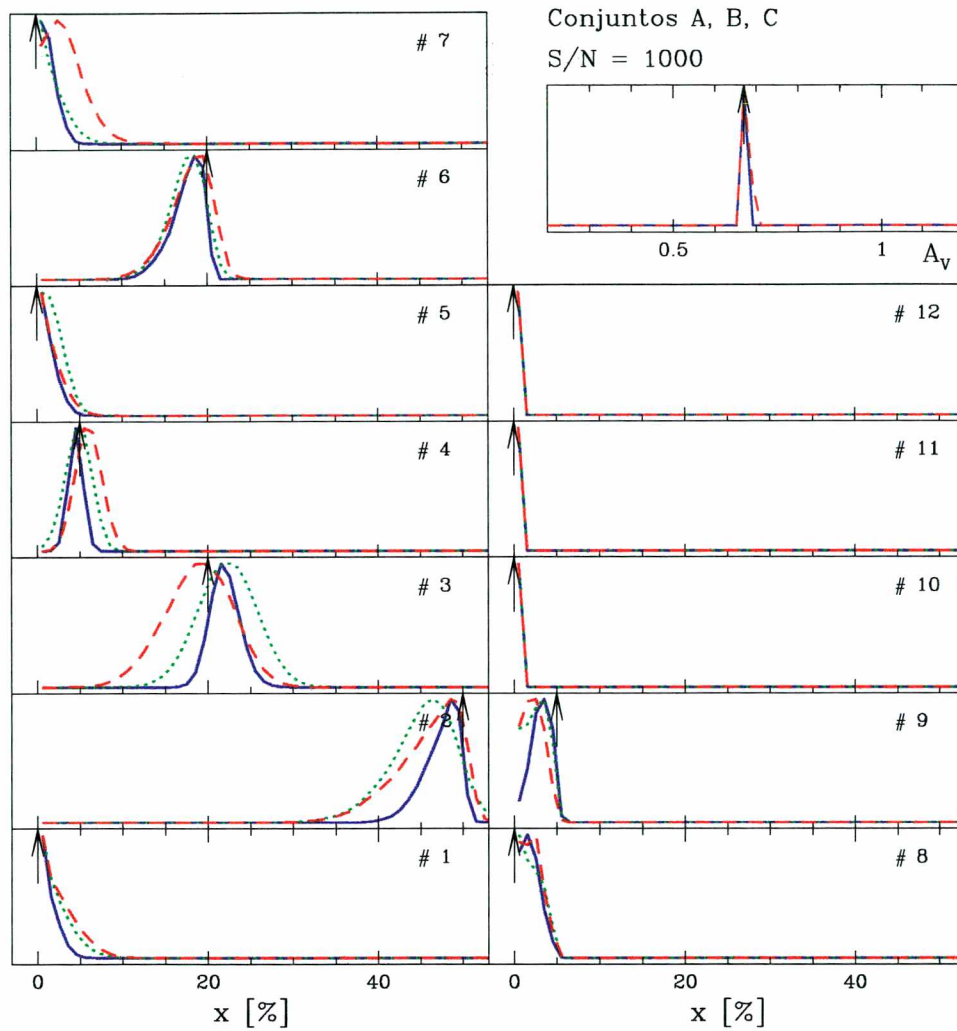


Fig. 4.10: O mesmo das figuras anteriores, mas a para a galáxia-fictícia 6 (com população relativamente velha). Seguimos as mesmas convenções de cores das 2 figuras anteriores.

resultados. Na figura 4.11 vemos um gráfico de $\overline{x_i}$ contra $x_{\text{verdadeiro}}$ para as 100 galáxias-fictícias e as quatro relações S/N indicadas. As idades estão indicadas horizontalmente a direita dos painéis e foi utilizado o grupo de observáveis A. Podemos verificar o ganho em precisão se compararmos os painéis para $S/N = 60$ com a figura 4.6 que utiliza a mesma qualidade de espectro. O mais notável nesta figura é a possibilidade de se obter bons resultados mesmo com espectros realísticos: $S/N \sim 60$. Alguns desvios ainda permanecem para as maiores idades e para espectros de pouca qualidade. A figura 4.12 mostra o mesmo resultado para os grupos B e C para os quais os resultados são igualmente melhorados.

4.4 A degenerescência entre idade e metalicidade

Mesmo para os resultados agrupados em idade das figuras 4.11 e 4.12 podemos notar que para $S/N \sim 10$ e 30 existe um efeito de compensação entre os grupos de idade mais velhos, 10^9 e 10^{10} anos. Isto se deve à degenerescência entre idade e metalicidade (O' Connell 1986 e 1994 e Worthey 1994) presente também na base de Bica e que afeta justamente as componentes mais antigas, devido à pouca resolução da base e sua consequente incapacidade em diferenciar se o espectro de determinada população é vermelha porque é mais metálica ou porque é mais velha. A figura 4.13 ilustra este efeito. Em seu eixo horizontal está representada a diferença do logaritmo das idades sintética e observada contra a diferença do logaritmo das metalicidades sintética e observada. Em cada quadro estão representadas diversas relações S/N e grupos de observáveis. Se $\Delta \log t > 0$, implica que se super-estimou a idade e portanto temos cores mais avermelhadas e linhas metálicas mais pronunciadas. Na média, isto tende a ser compensado pela metalicidade sub-estimada, que produz linhas mais azuis e linhas mais fracas. Fica claro, a partir desta figura, que a utilização de espectros de melhor qualidade, conjuntos mais completos de observáveis e a combinação de elementos de mesma idade podem quebrar a degenerescência entre idade e metalicidade, dentro dos limites de resolução da base.

4.5 Resultados agrupados em metalicidade

Dispomos de uma base caracterizada por dois parâmetros: idade e metalicidade, e vimos na seção 4.3 que os resultados agrupados em idade produzem resultados muito melhores. A figura 4.14 mostra um diagrama $x_{\text{verdadeiro}}$ contra x_{medio} , onde os resultados estão agrupados em metalicidade. Utilizou-se o conjunto A e as duas relações S/N indicadas, mas não se conseguiu melhorias nos resultados. Isto sugere que esta base não pode nos fornecer pistas precisas a respeito do enriquecimento químico dos objetos. Em Bica (1988) este problema foi minimizado, através da escolha de caminhos evolutivos no plano t vs. Z , que só permitiam o aumento da metalicidade. Neste trabalho entretanto não impusemos quaisquer destes vínculos, pois este grau de liberdade permite que a síntese responda por uma variedade maior de caminhos evolutivos.

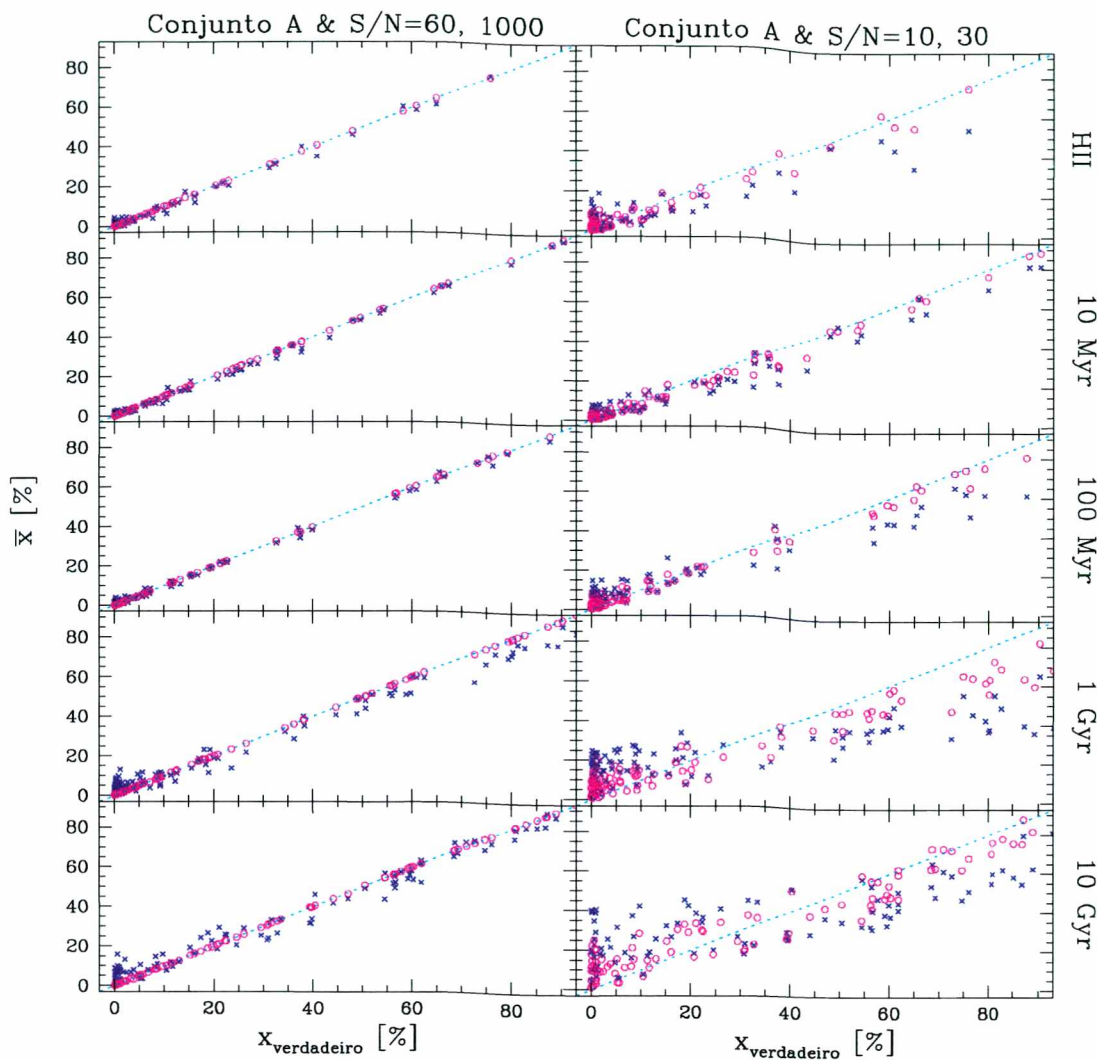


Fig. 4.11: Resultados agrupados em idade (mostradas à direita dos painéis) para o conjunto A e as relações sinal/ruído indicadas. Notar o enorme ganho em precisão, mesmo para espectros realísticos ($S/N \sim 30$). Comparar com a figura 4.6 que utiliza a mesma qualidade de espectro.

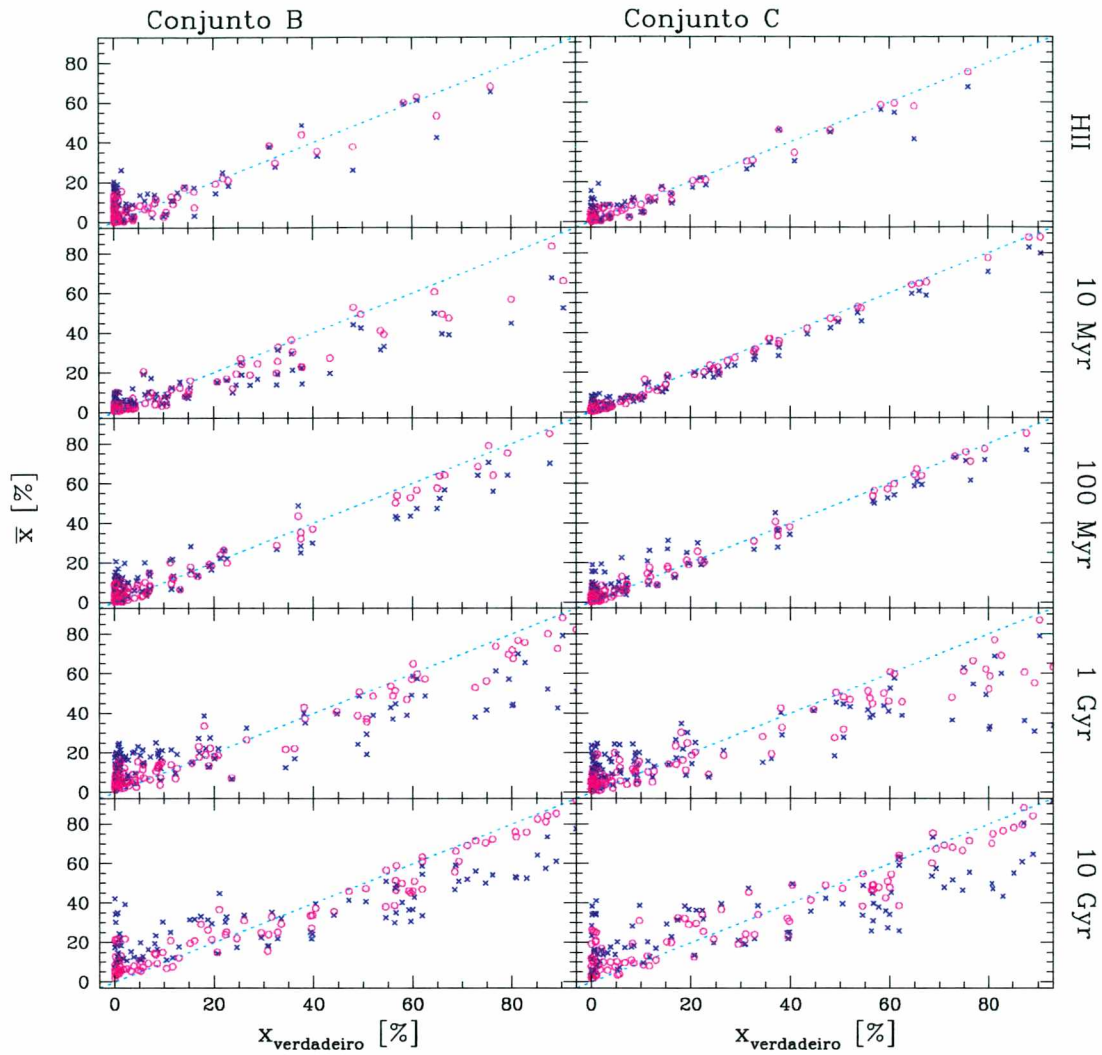


Fig. 4.12: O mesmo da figura anterior mas para os conjuntos B e C. Os círculos correspondem a $S/N = 100$ e as cruzes a $S/N = 30$.

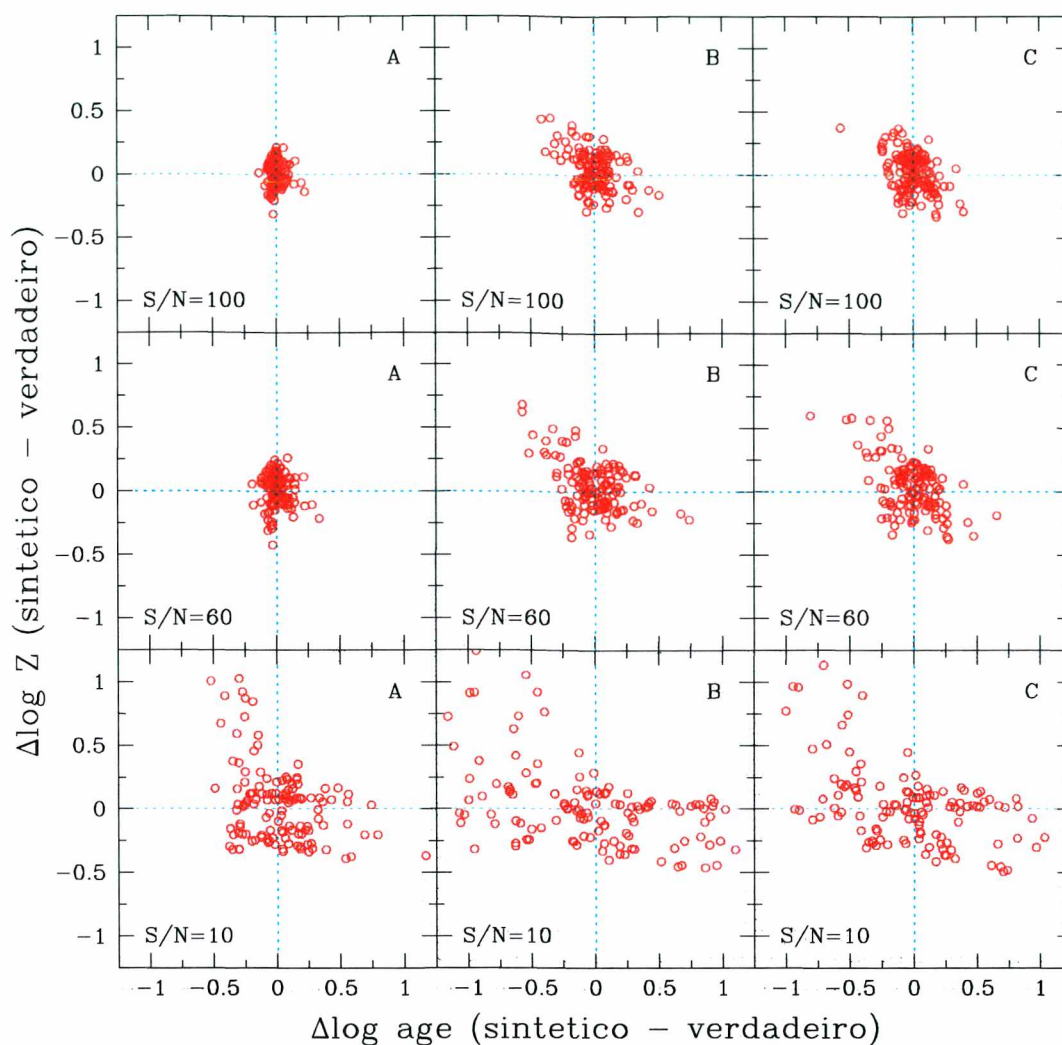


Fig. 4.13: Diferença de idades contra diferença de metalicidades (em escala logarítmica). Em cada painel estão indicados o conjunto de observáveis e a qualidade dos dados. O espalhamento é maior para espectros piores e para conjuntos de observáveis mais pobres. Notar como (na média) as idades são super-estimadas são compensadas por metalicidades sub-estimadas (e vice-versa).

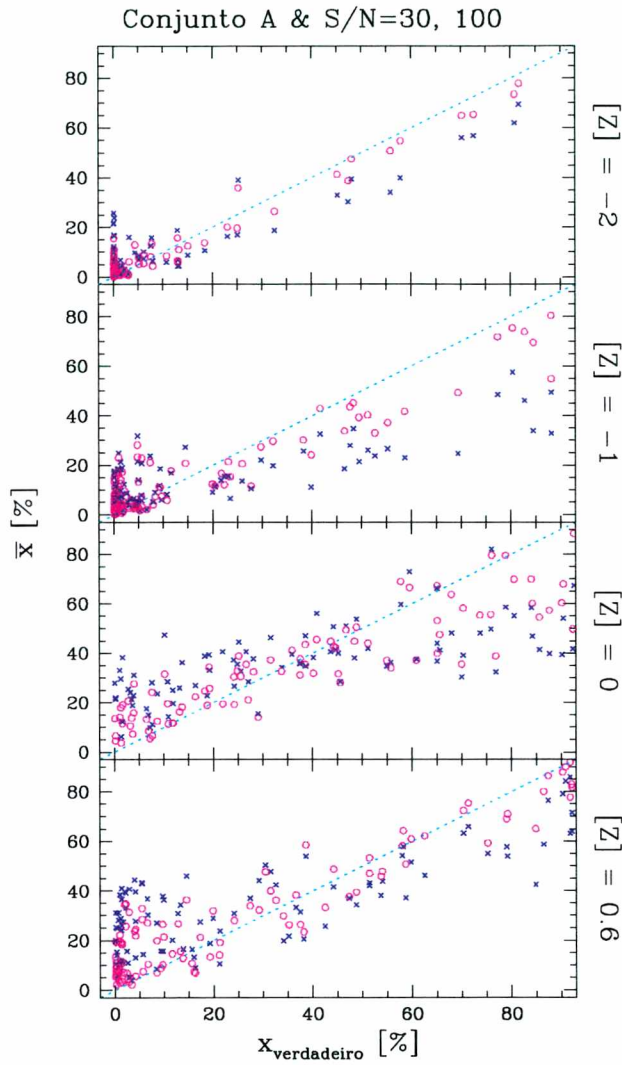


Fig. 4.14: Resultados combinados em metalicidade. Ao contrário do que se espera para os resultados combinados em idade, não se obtém resultados muito melhores.

4.6 Sumário dos resultados

Ao longo deste capítulo vimos como os resultados da síntese são afetados pela qualidade do espectro e pelo conjunto de observáveis utilizado. Foi possível verificar que tanto os erros quanto conjuntos pobres em observáveis tendem a aumentar as incertezas dos resultados e possuem essencialmente o mesmo efeito de redistribuir as curvas de distribuição para os parâmetros. Entretanto, as redistribuições não acontecem de forma aleatória, mas obedecem essencialmente a dois efeitos: a redistribuição entre elementos de mesma idade e à degenerescência entre idade e metalicidade. O primeiro efeito, a princípio negativo, nos conduziu à idéia de combinar os resultados em idade. Os resultados da seção 4.3 mostram o quanto esta maneira de apresentar resultados melhorou a síntese, mesmo para espectros realísticos como $S/N \sim 30$. Por outro lado, apresentar os resultados agrupados em metalicidade não produz melhores resultados e aponta a clara falta de poder de resolução da base com respeito à evolução química. Por fim, o efeito da degenerescência entre t e Z foi discutido na seção 4.4. Ao longo do capítulo mostramos os resultados em função do valor médio dos parâmetros, pois detectamos que \bar{x} é estatisticamente menos susceptível ao efeito dos erros que o melhor modelo (x_{melhor}).

Os capítulos 1 ao 4 estão presentes, numa forma menos detalhada, no trabalho Cid Fernandes *et al.* (2001). Nesta dissertação entretanto, faremos nos capítulos 5 e 6 algumas aplicações deste método.

Capítulo 5

Aplicação às galáxias do atlas de Kennicutt

No capítulo anterior testamos o algoritmo de Metropolis para galáxias fictícias e verificamos que pelo menos as idades das populações estelares podem ser recuperadas com precisão. Nesse capítulo vamos apresentar uma primeira aplicação do método a galáxias de verdade. Vamos utilizar 24 das galáxias do atlas espectro-fotométrico de Kennicutt (1992), que abrange 55 galáxias próximas de diversos tipos morfológicos.

Não faremos uma análise individual para todas as galáxias, mas pretendemos relacionar os resultados da síntese agrupados em idade com nosso conhecimento a respeito da relação entre morfologia de galáxias e suas populações estelares. Esperamos, portanto, que galáxias elípticas possuam boas frações de luz provenientes de populações velhas enquanto que as galáxias espirais possuam contribuições de população velha (devido ao bojo) e jovem (devido aos braços espirais).

5.1 Os dados

Para este estudo selecionamos 24 galáxias do atlas de Kennicutt (1992). Estas galáxias varrem praticamente toda a sequência de Hubble desde elípticas até irregulares (ver tabela 5.1). Nenhuma das galáxias apresenta evidências de peculiaridades, tais como núcleo ativo ou fusão. Todas as galáxias se encontram em altas latitudes Galácticas de modo que não sofrem grandes efeitos de extinção devido à Via Láctea. Os espectros cobrem o intervalo de 3650 até 7100 Å, com uma resolução de 5–8 Å. Os dados foram coletados no telescópio de 2.3m, no observatório de Steward e os espectros referem-se às galáxias inteiras, tendo sido utilizada a técnica de varredura conhecida como “drift scanning”, com uma fenda longa.

Como no capítulo anterior, os observáveis que sintetizaremos são larguras equivalentes e cores. As medidas destes observáveis foram feitas pelo Dr. H. Schmitt (NRAO) e seus respectivos valores estão mostrados na tabela 5.2.

Dentre os observáveis da base, optou-se por não sintetizar as linhas do CN e H β , pois a primeira é mais sujeita a erros e a segunda é frequentemente contaminada por emissão. Assim, sobram dentro da faixa observada as linhas do CaIIK, banda-G, MgI, H δ , e H γ . Além delas sintetizamos as cores C_{3660}/C_{5870} , C_{4020}/C_{5870} , C_{4510}/C_{5870} , C_{6630}/C_{5870} . Tipicamente os erros nas larguras equivalentes estão entre

As 24 galáxias do atlas de Kennicutt			
Galáxia	T	Tipo Espectral	Classificação de Hubble
NGC1357	1	-2.0	Sa
NGC1832	3	2.4	SBb
NGC2276	5	-	Sc
NGC2775	1	-3.4	Sa
NGC2903	5	1.1	Sc
NGC3147	3	0.0	Sb
NGC3245	-2	-4.6	S0
NGC3368	2	-3.3	Sab
NGC3379	-5	-5.2	E0
NGC3623	1	-4.1	Sa
NGC3627	3	1.4	Sb
NGC3941	0	-2.7	SB0/a
NGC4262	-2	-3.8	SB0
NGC4449	9	10.8	Sm/Im
NGC4472	-4	-5.7	E1/S0
NGC4631	5	9.4	Sc
NGC4648	-5	-4.4	E3
NGC4775	5	9.7	Sc
NGC4889	-5	-4.1	E4
NGC5248	4	1.4	Sdc
NGC5866	-2	-4.4	S0
NGC6181	5	3.1	Sc
NGC6217	4	4.8	SBbc
NGC6643	5	3.8	Sc

Tab. 5.1: Na segunda coluna, vemos a classificação morfológica T, na terceira o Tipo espectral, definido por Sodré & Cuevas (1997) e na quarta a classificação morfológica usual de Hubble.

0.5 e 1 Å. Por simplicidade fixamos os erros em todas as linhas em 0.5 Å, enquanto que os erros nas cores foram fixados em 0.05.

A figura 5.1 mostra os espectros para 4 das 24 galáxias que sintetizamos: NGC 3379, NGC 4262, NGC 1832 e NGC 4449. A partir do topo, os objetos estão ordenados em ordem decrescente de idade, como verificaremos na seção a seguir, e como é evidente por inspeção visual.

5.2 Resultados

Na seção 4.5 vimos que os resultados agrupados em metalicidade não apresentam grandes vantagens. Assim, na tabela 5.3 apresentamos somente os resultados da síntese agrupados em idade, acompanhados de seus respectivos erros. É interessante notar que em alguns casos, objetos com linhas de emissão muito fracas, i.e. apenas

Observáveis reais das 24 galáxias									
Galáxia	Larguras Equivalentes [Å]					Cores			
	CaIIK	Banda-G	MgI	H δ	H γ	3660	4020	4510	6630
NGC1357	11.8	7.3	7.8	3.8	5.5	0.507	0.711	0.956	0.939
NGC1832	7.7	5.5	5.5	5.4	5.1	0.857	1.081	1.115	0.878
NGC2276	4.2	4.9	5.4	5.2	2.9	1.211	1.428	1.291	0.872
NGC2775	15.7	9.0	8.6	3.9	5.0	0.482	0.697	0.913	0.891
NGC2903	6.9	5.0	6.2	5.8	5.7	0.733	0.938	1.006	0.888
NGC3147	9.2	6.5	8.9	3.3	5.5	0.610	0.867	1.055	0.931
NGC3245	18.2	9.9	8.1	4.1	5.4	0.453	0.612	0.858	0.895
NGC3368	14.7	8.8	7.1	5.0	5.9	0.478	0.673	0.897	0.917
NGC3379	17.7	10.4	9.0	4.1	5.8	0.380	0.590	0.870	0.980
NGC3623	15.7	9.1	8.5	4.7	5.5	0.411	0.612	0.882	0.907
NGC3627	8.9	5.7	5.6	6.5	6.3	0.739	0.950	1.039	0.901
NGC3941	16.7	10.0	9.7	4.1	6.0	0.455	0.695	1.025	0.882
NGC4262	18.4	10.2	10.3	5.8	5.6	0.436	0.691	0.908	0.876
NGC4449	3.7	3.7	3.8	4.7	-	1.225	1.773	1.485	0.818
NGC4472	16.5	8.9	9.4	5.3	5.6	0.358	0.541	0.831	0.925
NGC4631	3.4	3.4	4.9	5.6	4.2	1.553	1.648	1.421	0.865
NGC4648	18.1	9.3	9.2	4.7	5.4	0.380	0.611	0.883	0.861
NGC4775	4.1	3.9	4.2	5.7	3.8	1.353	1.659	1.398	0.794
NGC4889	16.8	10.1	11.3	3.6	5.9	0.387	0.634	0.945	0.890
NGC5248	7.9	3.9	4.2	5.7	3.8	0.793	0.923	1.019	0.879
NGC5866	18.4	10.2	10.3	5.8	5.6	0.384	0.616	0.861	0.989
NGC6181	6.0	4.5	5.1	5.6	5.0	0.901	1.169	1.081	0.846
NGC6217	8.3	4.8	4.2	6.1	6.0	0.988	1.300	1.228	0.888
NGC6643	6.8	3.0	5.7	5.1	4.7	0.942	1.178	1.221	0.877

Tab. 5.2: Observáveis calculados por H. Schmitt. Os fluxos estão normalizadas em 5870 Å.

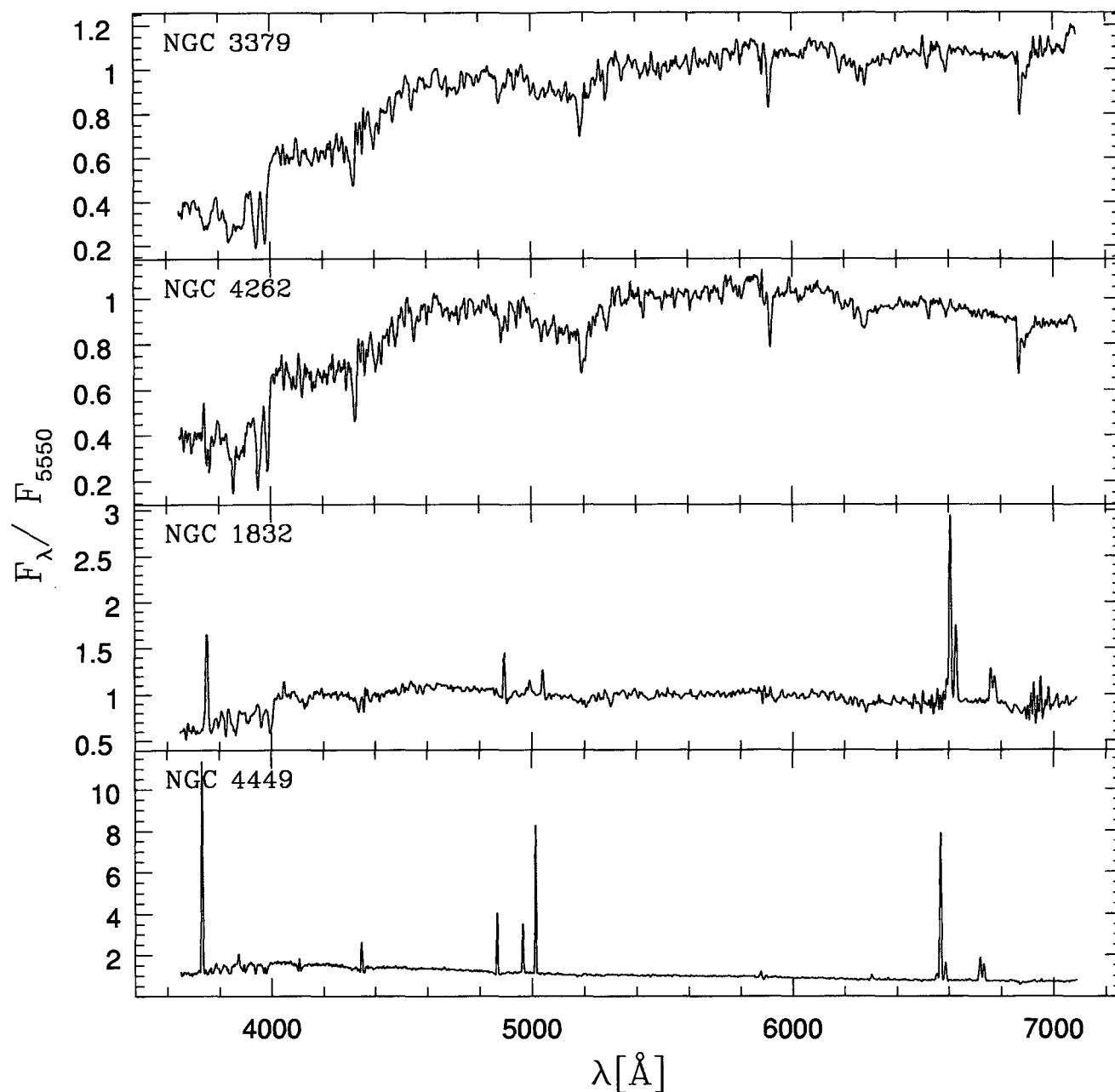


Fig. 5.1: Espectros integrados de três galáxias do atlas de Kennicutt. Os objetos estão ordenados a partir do topo em ordem decrescente de contribuição de populações velhas para sua luz.

Síntese agrupada em idade e erros					
Galáxia	10^{10} anos	10^9 anos	10^8 anos	10^7 anos	10^6 anos
NGC1357	76.1±4.2	13.0±4.1	2.0±1.3	2.9±2.1	5.9±1.5
NGC1832	37.3±5.9	34.8±6.2	7.0±2.7	6.7±4.0	14.2±2.3
NGC2276	20.4±4.4	32.4±5.3	10.0±3.0	6.6±4.0	30.6±2.3
NGC2275	80.5±3.7	14.9±3.5	0.4±0.4	1.2±1.0	2.9±0.9
NGC2903	50.3±6.9	17.1±6.6	13.2±3.0	10.5±5.6	8.8±3.0
NGC3147	65.5±3.8	15.4±4.1	4.3±1.8	2.8±2.1	12.0±1.4
NGC3245	89.7±3.1	9.3±3.1	0.2±0.3	0.5±0.5	0.2±0.3
NGC3368	77.1±4.1	18.9±3.9	1.0±0.8	1.8±1.3	1.2±0.9
NGC3379	89.6±3.0	8.8±3.0	0.4±0.4	0.8±0.6	0.4±0.4
NGC3623	82.9±3.6	13.4±3.6	0.9±0.7	1.7±1.2	1.2±0.8
NGC3627	65.5±3.8	15.4±4.1	4.3±1.8	2.8±2.1	12.0 ±1.4
NGC3941	74.0±3.7	22.6±3.6	0.4±0.4	0.9±0.8	2.1 ±0.8
NGC4262	70.6±3.5	28.1±3.6	0.3±0.3	0.6±0.5	0.3±0.3
NGC4449	8.4±3.2	23.2±4.4	34.8±3.8	2.8±2.1	30.8±2.0
NGC4472	83.6±3.5	13.0±3.6	1.3±0.9	1.6±1.0	0.6±0.5
NGC4631	5.1±2.4	29.4±3.8	20.6±3.1	8.2±4.5	36.7±2.5
NGC4648	84.6±3.3	14.0±3.3	0.4±0.4	0.7±0.6	0.4±0.4
NGC4775	7.1±3.0	32.4±4.2	22.0±3.2	5.0±3.3	33.4±2.1
NGC4889	85.0±3.2	11.5±3.3	0.5±0.5	1.0±0.9	2.0±0.7
NGC5248	40.6±7.4	31.1±7.3	7.1±2.9	13.2±6.0	8.0±3.2
NGC5866	77.8±3.3	20.9±3.4	0.4±0.4	0.7±0.5	0.2±0.3
NGC6181	36.2±6.5	25.1±6.5	13.1±3.3	10.7±5.6	14.9±3.0
NGC6217	15.8±4.9	51.4±6.0	11.5±3.2	6.1±3.8	15.1±2.1
NGC6643	32.3±5.9	28.1±6.0	12.7±3.1	8.6±4.9	18.2±2.7

Tab. 5.3: Resultados da síntese agrupados em idade para as 24 galáxias do atlas de Kennicutt. Os erros acompanham as respectivas abundâncias em %.

$H\alpha$ em emissão, a síntese indica uma componente fraca, mas não nula, de uma população tipo região HII. Ex: NGC 1357 e NGC 3827. Mostramos ainda na tabela 5.4 os observáveis sintéticos para as 24 galáxias estudadas. É interessante comparar o conteúdo desta tabela com aquele da tabela 5.2.

As figuras 5.2 a 5.5 mostram os posteriores combinados em idade para as 4 galáxias cujo espectro apresentamos na seção anterior. Nestas figuras, os grupos de idade estão indicados verticalmente e a fração de luz (em 5870 Å) correspondente na horizontal. Como vemos, NGC 3379 é uma galáxia elíptica e a síntese indica justamente a predominância de populações estelares velhas. A galáxia NGC 4262 (tipo SB0) é dominada por estrelas velhas devido ao seu bojo proeminente. Uma população significativa de estrelas de 10^9 anos aparece, provavelmente devido a seu disco. As populações jovens são significativas nas galáxias NGC 1832 (SBb) e principalmente NGC 4449 (Sm/Im). Note que ambas possuem linhas de emissão oriundas de regiões HII, bem como linhas de Balmer em absorção de alta ordem

Observáveis Sintéticos das 24 galáxias									
Galáxia	Larguras Equivalentes [Å]					Cores			
	CaIIK	Banda-G	MgI	H δ	H γ	3660	4020	4510	6630
NGC1357	11.7	7.1	7.2	4.9	4.8	0.542	0.697	0.882	0.967
NGC1832	8.0	4.7	5.2	5.7	5.1	0.875	1.084	1.105	0.914
NGC2276	5.5	3.2	4.1	4.4	4.0	1.240	1.377	1.264	0.872
NGC2775	15.0	8.6	8.1	5.3	5.2	0.451	0.613	0.832	0.978
NGC2903	7.5	4.5	5.5	6.0	5.5	0.744	0.971	1.052	0.932
NGC3147	9.6	6.1	7.1	4.7	4.5	0.673	0.813	0.949	0.951
NGC3245	17.2	9.4	8.3	5.3	5.3	0.372	0.534	0.787	0.989
NGC3368	14.4	8.1	7.3	5.9	5.5	0.464	0.660	0.858	0.972
NGC3379	17.3	9.6	8.8	5.3	5.3	0.357	0.513	0.777	0.993
NGC3623	15.4	8.7	8.2	5.6	5.4	0.414	0.585	0.817	0.983
NGC3627	9.6	6.1	7.1	4.7	4.5	0.673	0.813	0.949	0.951
NGC3941	16.2	9.0	8.8	6.1	5.6	0.437	0.617	0.837	0.978
NGC4262	17.7	9.6	9.2	6.8	6.0	0.404	0.603	0.832	0.980
NGC4449	4.2	2.1	3.1	5.8	5.3	1.335	1.643	1.415	0.835
NGC4472	16.2	9.1	8.7	5.8	5.5	0.376	0.545	0.803	0.988
NGC4631	4.3	2.3	3.2	4.7	4.2	1.463	1.636	1.400	0.839
NGC4648	17.3	9.5	8.7	5.7	5.5	0.375	0.544	0.795	0.988
NGC4775	4.7	2.5	3.4	5.1	4.6	1.370	1.586	1.380	0.843
NGC4889	16.6	9.5	9.2	5.3	5.2	0.382	0.529	0.788	0.991
NGC5248	8.3	4.8	5.4	6.1	5.4	0.772	0.990	1.058	0.931
NGC5866	17.9	9.8	9.3	6.4	5.8	0.372	0.553	0.806	0.987
NGC6181	6.5	3.8	4.6	5.7	5.1	0.928	1.149	1.142	0.907
NGC6217	7.8	4.3	4.7	6.6	5.7	0.988	1.267	1.200	0.890
NGC6643	6.6	3.8	4.7	5.4	4.9	0.985	1.194	1.168	0.899

Tab. 5.4: Observáveis Sintéticos das 24 galáxias do atlas de Kennicutt. Comparar esta tabela com os dados constantes na tabela 5.2.

visíveis na região de 4000 Å dos espectros. Esta componente estelar jovem é bem detectada pela síntese.

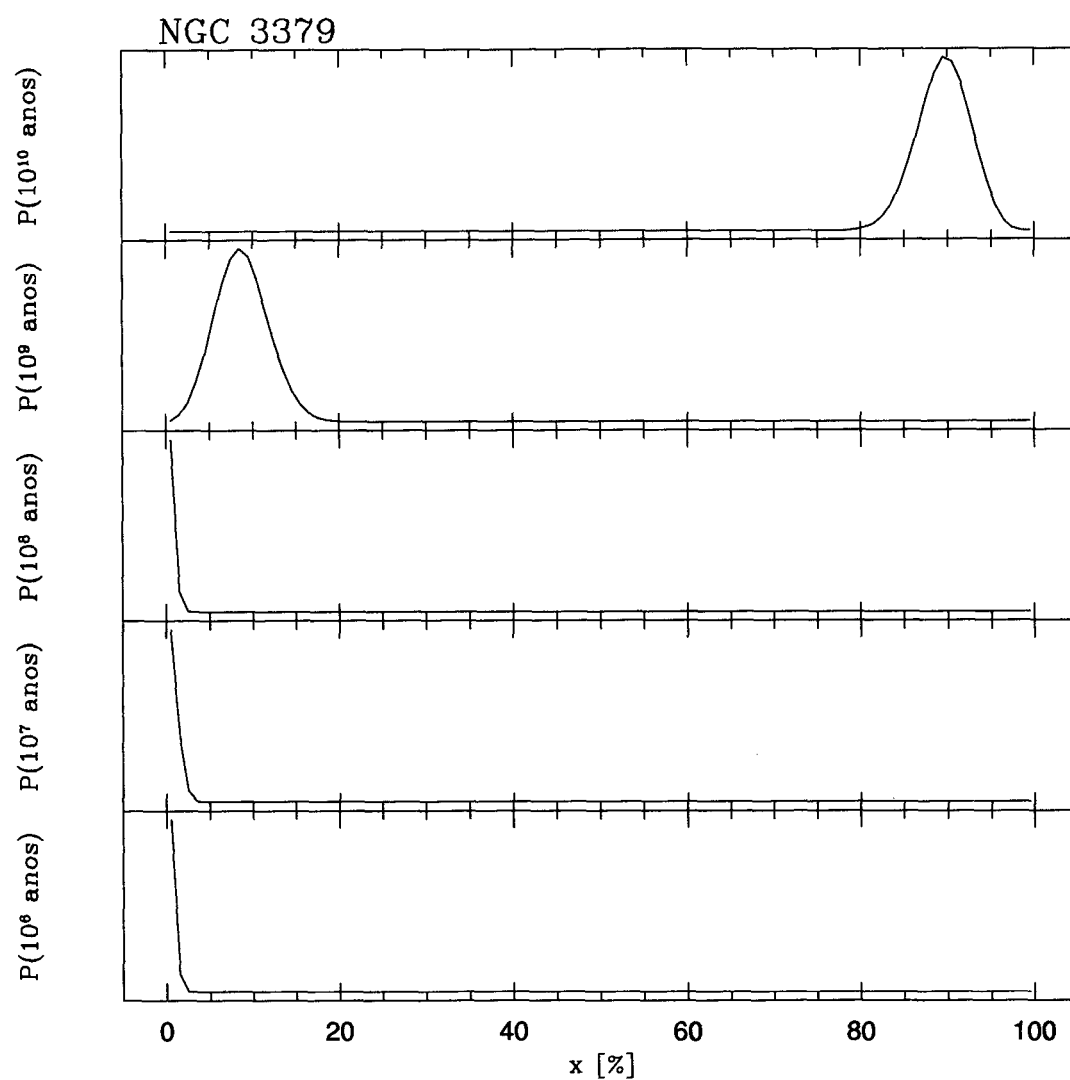


Fig. 5.2: Posteriores combinados em idade para a galáxia NGC3379 (elíptica). Como deveríamos esperar, vemos uma forte contribuição de populações velhas para a luz deste objeto.

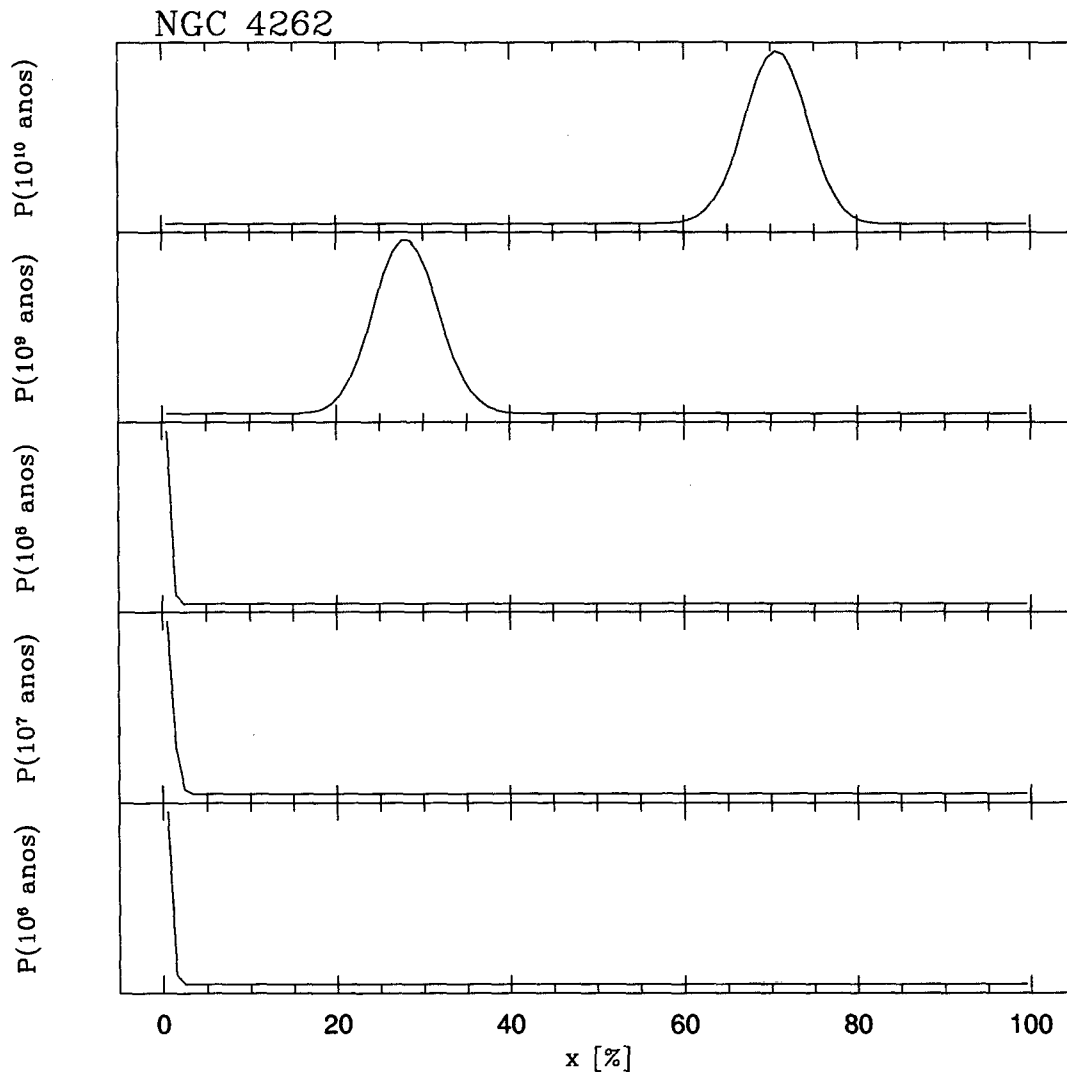


Fig. 5.3: Posteriores combinados em idade para a galáxia elíptica NGC4262. Notar que em relação à galáxia anterior existe uma contribuição mais forte da componente de idade $t = 10^9$ anos.

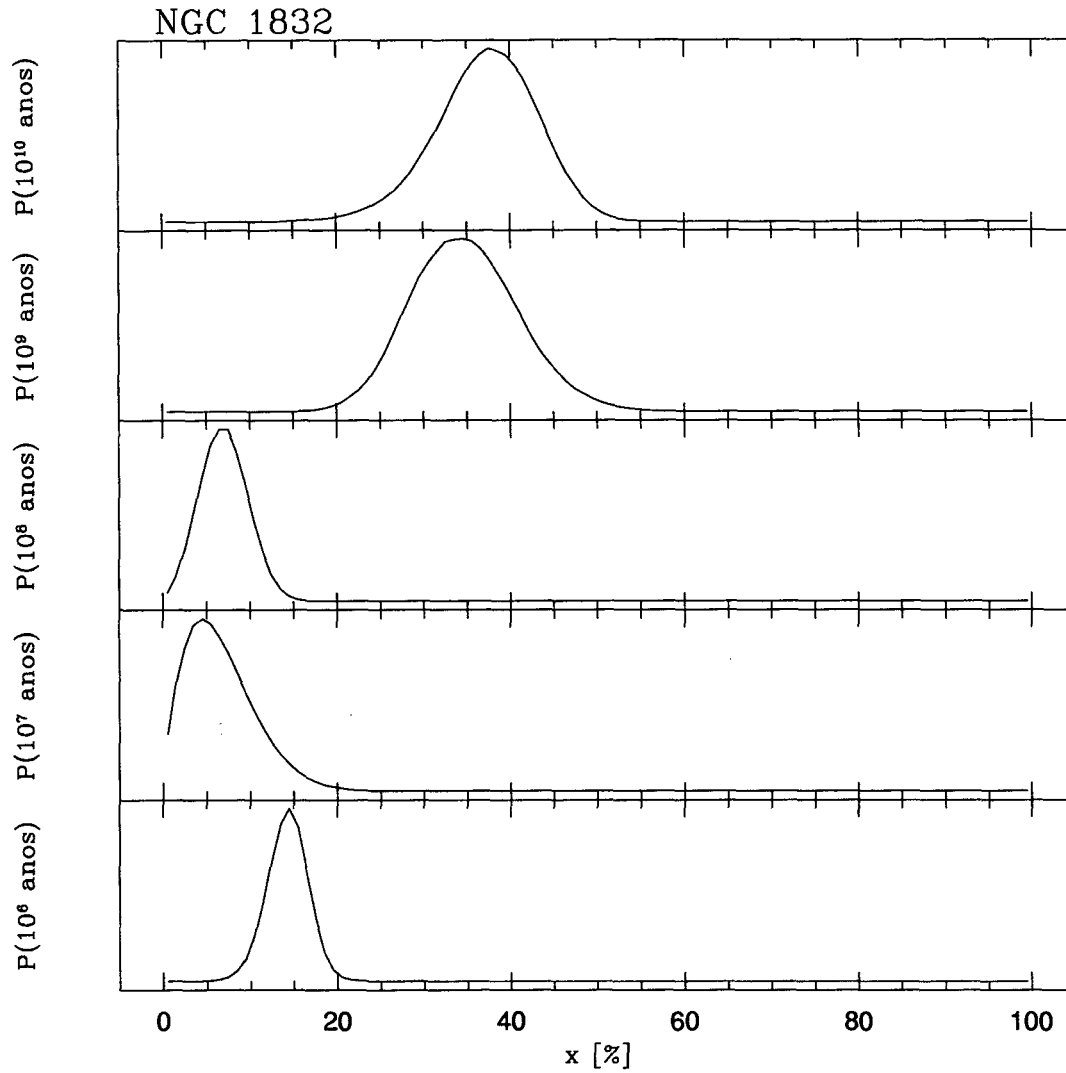


Fig. 5.4: Posteriores combinados em idade para a galáxia NGC1832. Parte da luz desta galáxia provém de populações jovens e de regiões HII, como esperaríamos para uma galáxia espiral.

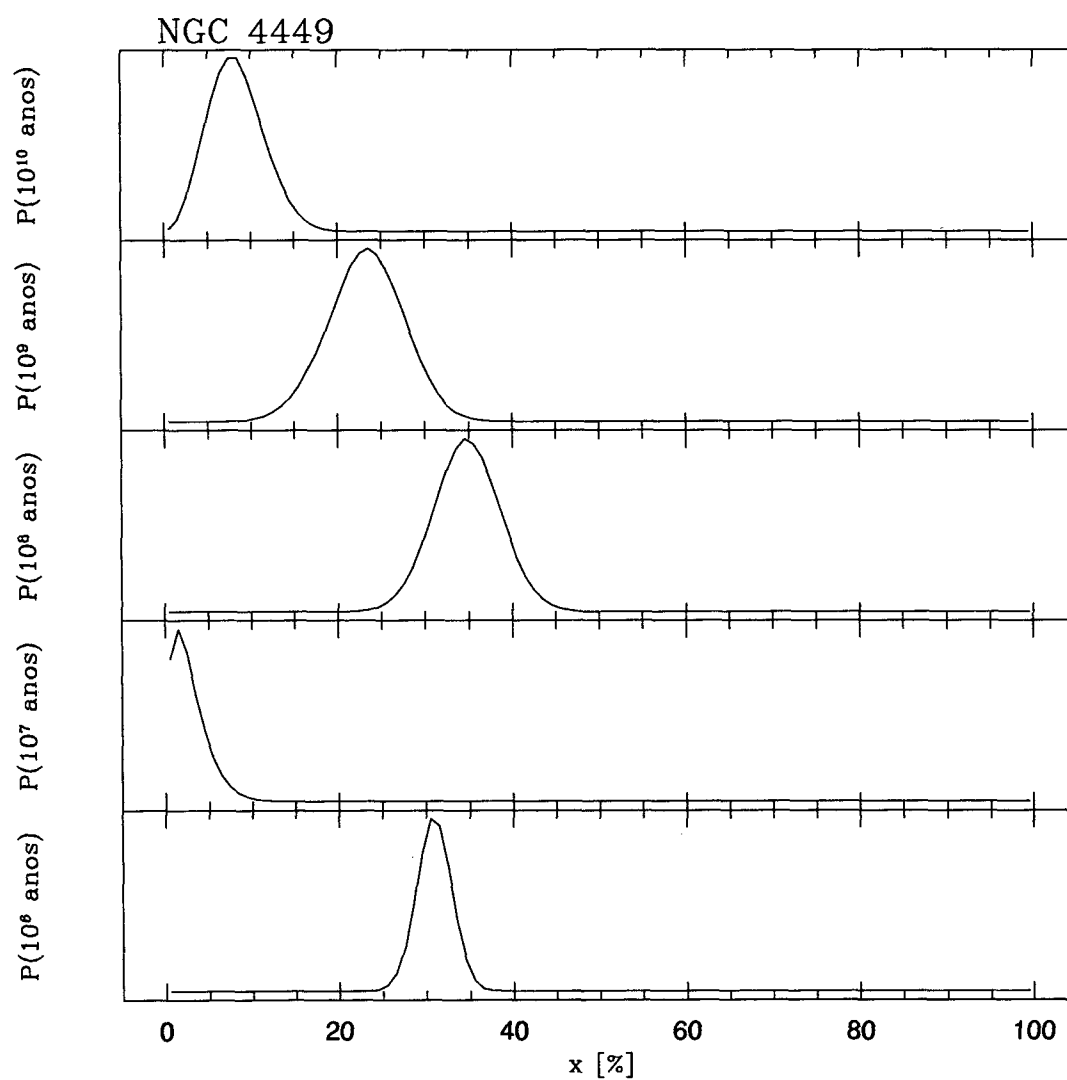


Fig. 5.5: Posteriores para a galáxia NGC 4449 (Sm/Im). Boa parte da luz deste objeto provém de populações jovens e de regiões HII, certamente devido ao seu disco.

5.3 Conclusões

Não objetivamos a discussão dos resultados da síntese para cada galáxia individualmente e nem uma interpretação detalhada dos resultados. É interessante porém investigar se existe alguma relação entre a população estelar, inferida pela síntese, e o tipo morfológico das galáxias.

Sabe-se, por exemplo que a sequência de Hubble para galáxias espirais corresponde a uma progressão da proporção entre as componentes do bojo e do disco; tal que espirais do tipo prematura (S0 – Sa) possuem um bojo mais pronunciado que o disco, enquanto as do tipo tardio (Sc – Sm/Im) possuem o disco proeminente e um bojo compacto. Sabemos também que o bojo possui populações mais antigas semelhantes àsquelas encontradas em galáxias elípticas, enquanto que os discos espirais possuem misturas de populações velha e nova, incluindo estrelas que ainda estão se formando, com regiões HII associadas a elas.

Dadas estas características, esperamos que populações jovens sejam progressivamente mais importantes para galáxias do tipo tardio, e o oposto para galáxias do tipo prematuro. As figuras 5.6 e 5.7 confirmam esta expectativa.

Na figura 5.6 estão mostradas as cinco componentes de idade contra o tipo morfológico T . Os valores de T neste gráfico vão desde $T = -5$, correspondendo a elípticas E0–E4, até $T = 9$, correspondendo a espirais Sm/Im. Assim, o tipo morfológico T indica, em ordem crescente, a importância das populações jovens para a luz dos objetos. Isto pode ser visualizado melhor na figura 5.7 onde reduzimos ainda mais a resolução em idade da base: combinamos as idades 10^9 e 10^{10} anos numa componente x_{velha} e as componentes 10^6 , 10^7 e 10^8 numa componente x_{nova} . Podemos ver claramente confirmadas as expectativas: a medida que cresce o valor de T , as populações velhas caem (painel superior) e as populações jovens (painel inferior) ganham importância, numa relação praticamente de 1 para 1.

Outra comparação interessante que podemos fazer é como o resultado da síntese se relaciona com outras técnicas de análise de espectros. Uma destas técnicas que vem sendo explorada para análise automática de espectros é a ACP (Análise de Componentes Principais). Sodré e Cuevas (1997) usam esta técnica para definir um índice a partir da projeção de um espectro num eixo principal. Este índice, denominado tipo espectral, usa mais informações (todo o espectro), resumindo-o a um número. O tipo espectral (tabela 5.1) foi calculado por L. Sodré para as galáxias do atlas de Kennicutt e na figura 5.8 mostramos a relação entre as componentes x_{velha} e x_{nova} contra o tipo espectral. Mais uma vez vemos uma clara correspondência de praticamente 1 para 1.

Estes resultados demonstram uma consistência gratificante do método de síntese aqui descrito, que resume o espectro, e o de ACP, que utiliza mais informações.

Outro fator importante é que os resultados da síntese com o método aqui descrito é capaz de recuperar resultados conhecidos a respeito da morfologia de galáxias. Isto nos faz pensar em diversas outras aplicações futuras deste método, conforme discutiremos na seção 7.1.

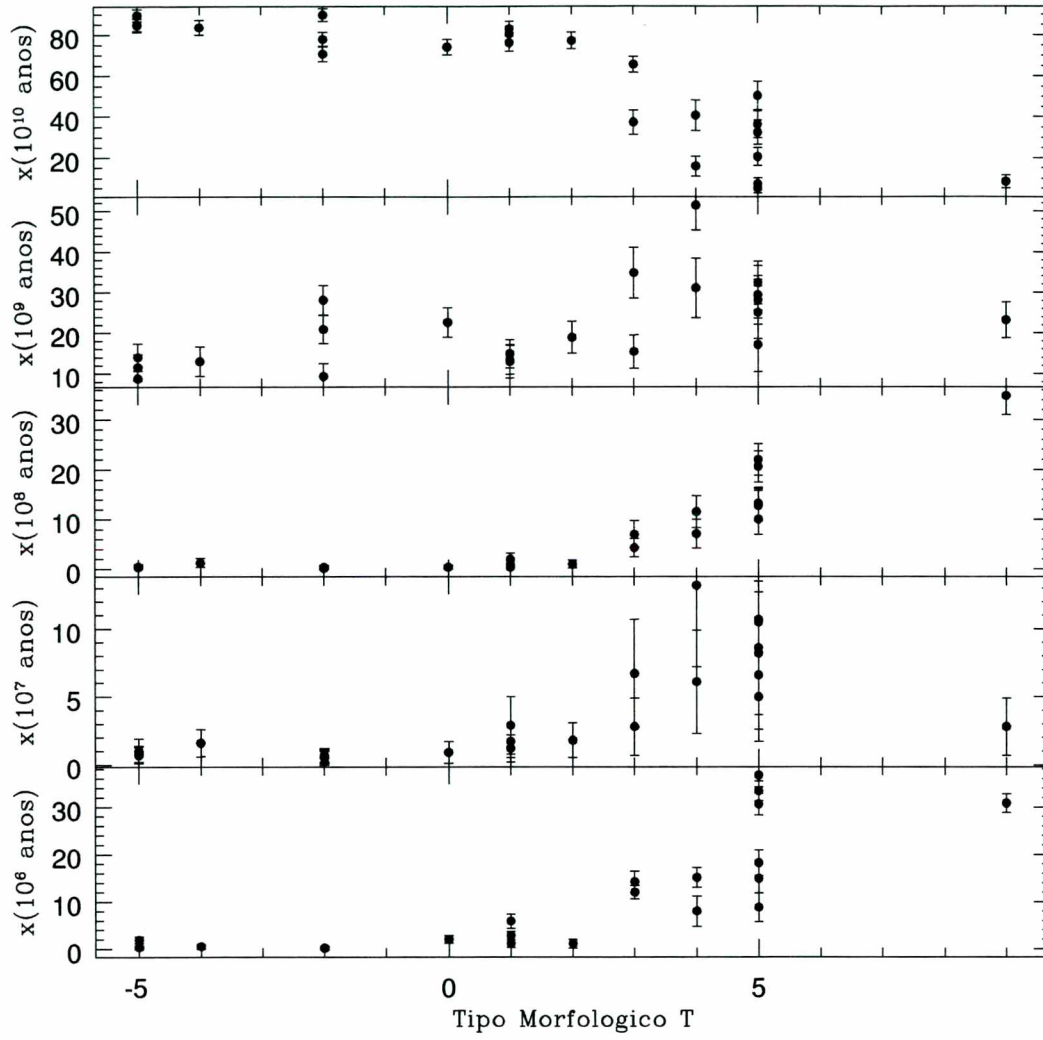


Fig. 5.6: Tipo morfológico T em função da idade. Notar que a medida que o valor de T cresce, aumenta a contribuições das populações jovens para a luz dos objetos.

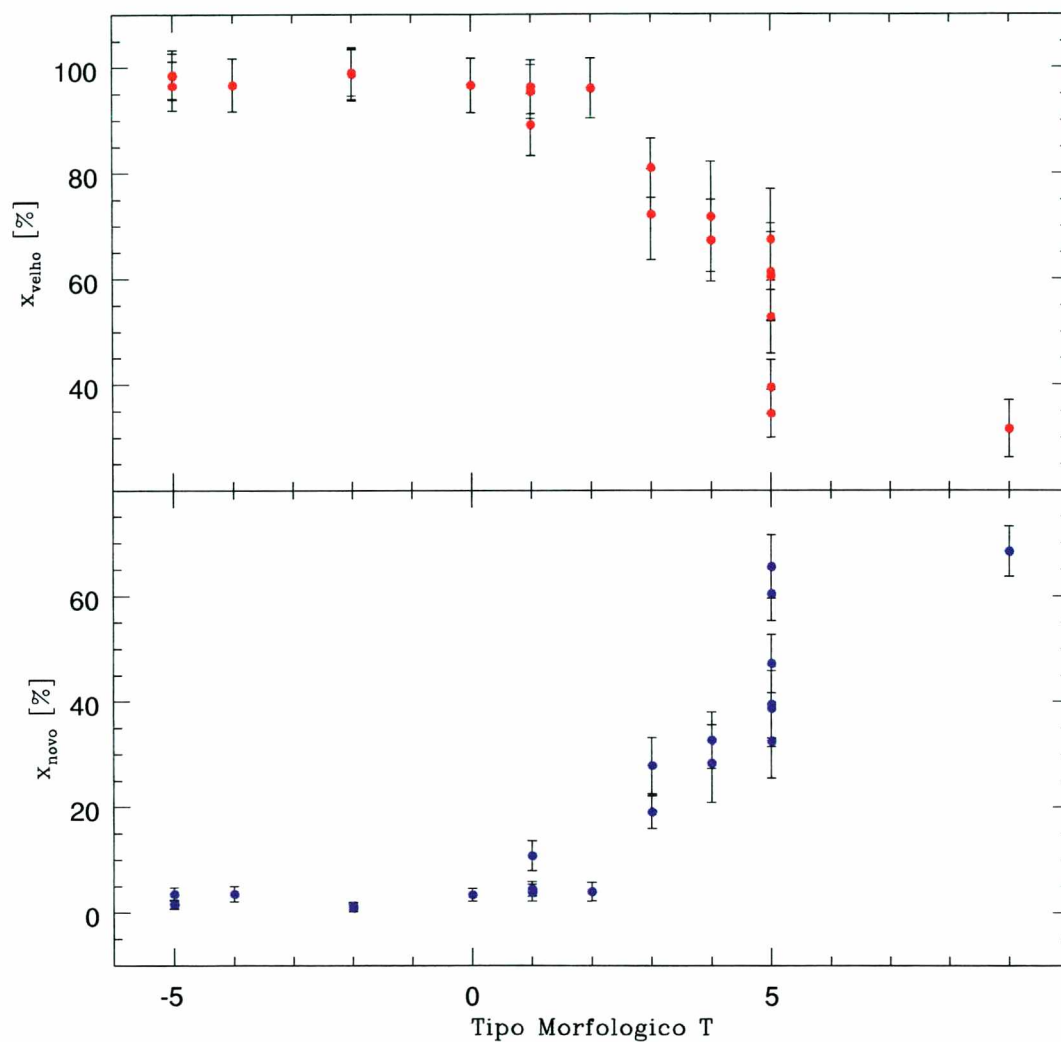


Fig. 5.7: O mesmo da figura anterior, mas para as idades combinadas em apenas duas: $x_{velha} = 10^9$ e 10^{10} anos e $x_{nova} = 10^8, 10^7$ e 10^6 anos.

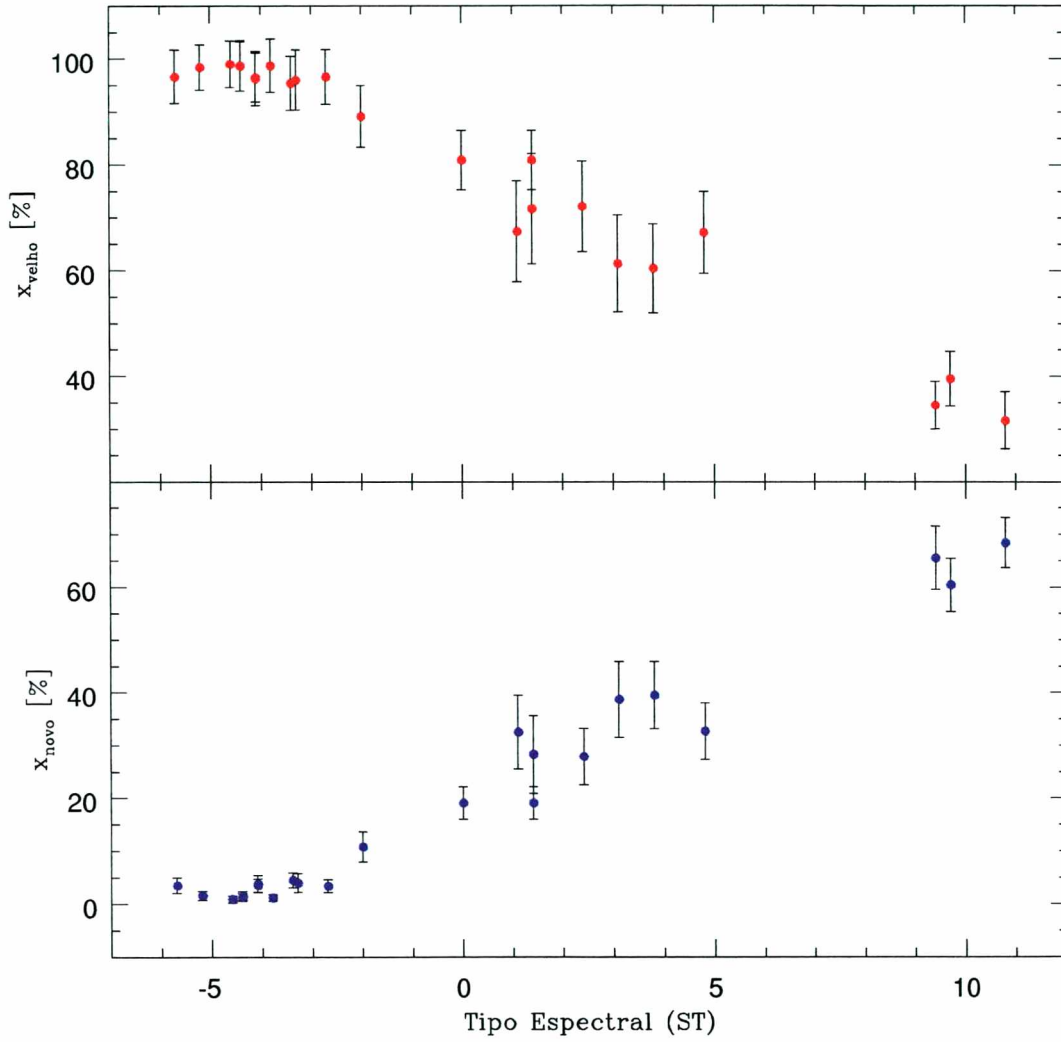


Fig. 5.8: Variação do tipo espectral em função da idade para as componentes x_{velha} e x_{nova} . Vemos que o crescimento do tipo espectral indica o aumento da contribuição das populações jovens para a luz do objeto.

Capítulo 6

O Domínio Sintético da base

Neste capítulo vamos apresentar uma aplicação do conceito de **domínio sintético** à base de 12 elementos que usamos ao longo desse trabalho. Os resultados são bastante ilustrativos no que diz respeito ao problema da **degenerescência algébrica** do problema de SPESE com essa base. Também a título de ilustração, investigamos a posição de 29 galáxias do atlas de Kennicutt ¹ em relação ao domínio sintético em uma série de diagramas. Conclui-se que na prática a não unicidade matemática do problema de SPESE não é um problema grave. A “degenerescência estatística” que resulta da presença de erros observacionais é mais relevante no processo de síntese, pelo menos no que se refere a essa base.

6.1 O conceito de Domínio Sintético

O conceito de domínio sintético (DS) foi elaborado por Pelat (1997 e 1998), em sua elegante formulação matemática do problema de SPESE. O DS simplesmente mapeia a região no espaço de observáveis correspondente a todas possíveis combinações dos elementos da base.

Nosso objetivo aqui é possibilitar uma visualização do domínio sintético, de modo que se possa avaliar de maneira gráfica se uma dada galáxia é “**algebricamente sintetizável**” pela base ou não. Por “algebricamente sintetizável” entende-se que seus observáveis podem ser reproduzidos **exatamente** por combinações da base. Objetos fora do domínio sintético portanto não podem ser sintetizados matematicamente de forma exata, e o problema de degenerescência algébrica (não unicidade) não se aplica.

O espaço de observáveis na base de Bica (1988) tem 16 dimensões, sendo 9 de W 's e 7 de cores. Para visualizar o domínio sintético nesse espaço somos forçados a apresentar apenas projeções bi-dimensionais. Esta não é uma limitação séria, pois como veremos, tais cortes já são bastante úteis para julgar a possibilidade de sintetizar galáxias.

Para um par de larguras equivalentes, digamos W_1 e W_2 , o domínio sintético no plano W_1 vs W_2 é definido por um conjunto de equações que definem estas W_i em termos de **pares** de populações, isto é, em termos de apenas 2 elementos da base

¹ Todas as que sintetizamos e outras cinco

(ver Pelat 1997, 1998). Os valores de W_1 que podem ser sintetizados apenas com os elementos j e k da base são:

$$W_1 = \frac{x_j W_{1j}^* F_{1j}^* + x_k W_{1k}^* F_{1k}^*}{x_j F_{1j}^* + x_k F_{1k}^*} \quad (6.1)$$

onde, pelo vínculo de normalização, $x_j + x_k = 1$, já que todos outros elementos de \vec{x} são, por construção, nulos. A equação correspondente para W_2 é

$$W_2 = \frac{x_j W_{2j}^* F_{2j}^* + x_k W_{2k}^* F_{2k}^*}{x_j F_{2j}^* + x_k F_{2k}^*} \quad (6.2)$$

Deste modo, as populações j e k definem uma curva no plano W_1 vs W_2 (Pelat 1998). Ao longo desta curva o par (W_1, W_2) percorre o trajeto de (W_{1j}^*, W_{2j}^*) até (W_{1k}^*, W_{2k}^*) a medida que x_j varia de 1 a 0 (e portanto x_k de 0 a 1). Permutando sobre todos pares possíveis de $j = 1 \dots 12$ e $k = 1 \dots 12$ da base obtém-se um “emaranhado” de $12 \times (12 - 1)/2 = 66$ curvas, cuja **envoltória** define o domínio sintético no plano W_1 vs W_2 .

6.2 Visualização do Domínio sintético para a base de Schmidt et al. (1991)

Exemplos didáticos do domínio sintético com bases de poucos elementos foram apresentados por Pelat (1998). Na Figura 6.1 apresentamos o DS projetado no plano W_K vs W_{MgI} para a base de Schmidt et al (1991). O importante nesta e nas demais figuras é a envoltória definida pelas 66 curvas graficadas. Os números 1 à 12 nos gráficos correspondem à posição dos respectivos elementos da base (ver tabela 2.1). Como esperado, os elementos 1 (população mais velha e de maior metalicidade) e 12 (população mais jovem) representam sempre extremos nestes gráficos.

Note como a região sintetizável é bem mais restrita que o retângulo definido pelos máximos e mínimos de W_K e W_{MgI} , ocupados pelas populações 1 e 12 respectivamente. Este retângulo, chamado de “**retângulo sintético**” por Pelat, define limites óbvios da base: Não é possível sintetizar W 's maiores que a máxima ou menores que a menor da base. A maior parte da área dentro do retângulo sintético, porém, tão pouco pode ser sintetizada com a base, o que ilustra a utilidade do conceito de DS introduzido por Pelat.

6.2.1 O Domínio Sintético em diagramas W_i vs. W_i

Nas Figuras 6.2 e 6.3 apresentamos o DS projetado em vários outros planos de W 's. Entendemos que as figuras acima são bastante úteis na análise de dados reais, pois a simples posição de uma galáxia nestes diagramas pode de antemão definir se ela pode ou não ser sintetizada pela base. Caso seus observáveis se enquadrem dentro das regiões delimitadas nos gráficos acima, a galáxia é exatamente sintetizável, e a síntese é algebricamente degenerada. Caso ela se localize fora de qualquer desses diagramas, a base não é capaz de sintetizá-la exatamente, e **não existe o problema de não unicidade algébrica**, pelo simples fato de que não existe solução algébrica

que satisfaça todos os vínculos do problema. Note que esta conclusão independe de quantos observáveis são utilizados na síntese.

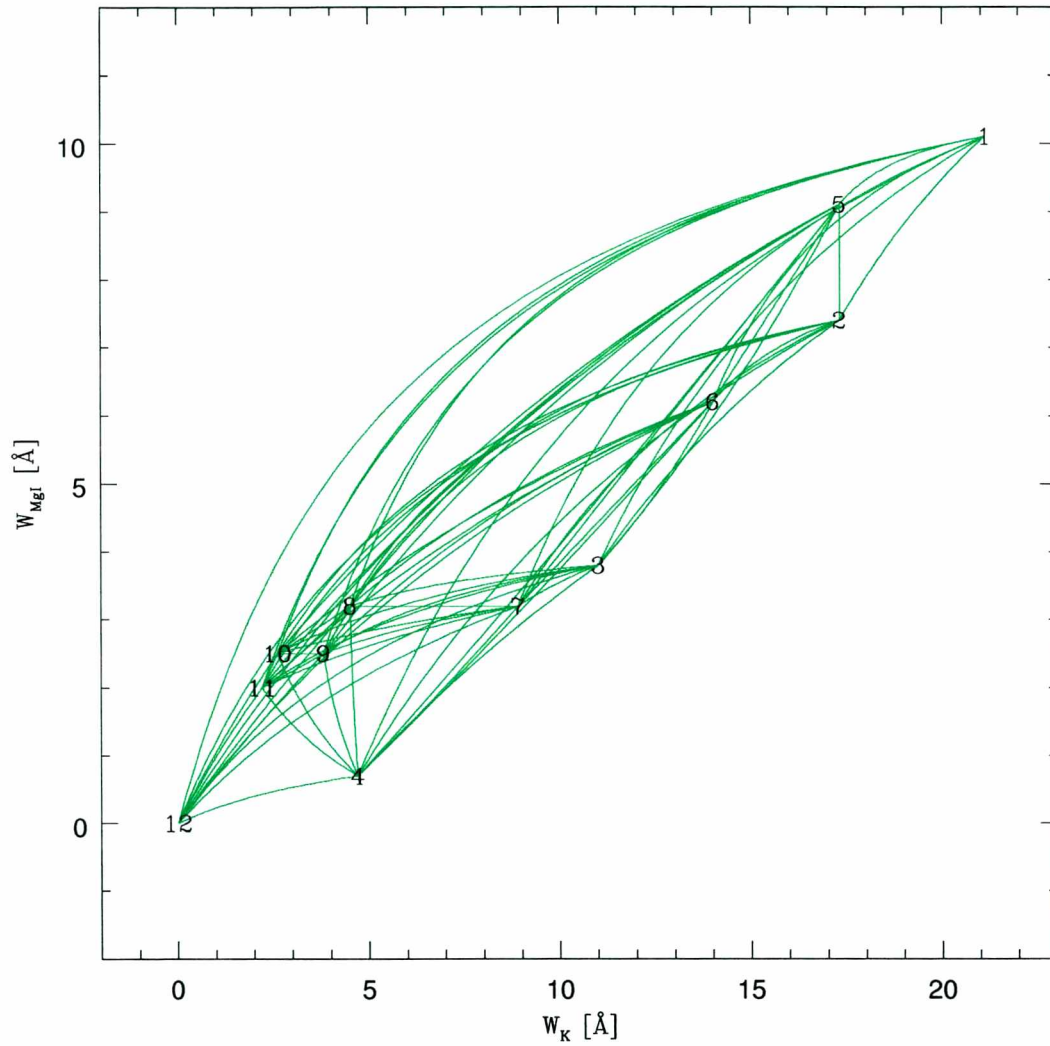


Fig. 6.1: Projeção do domínio sintético no plano definido pelas W 's das linhas CaII K e MgI. Os números indicam a localização dos 12 elementos da base. As linhas conectando cada par de elementos indicam as W 's que resultam de combinações apenas deste par de elementos. Galáxias com W 's fora da envoltória definida por este emaranhado de linhas não podem ser modeladas pela base, ou seja, não existe nenhuma combinação fisicamente aceitável de elementos da base que reproduza de forma exata as W 's medidas.

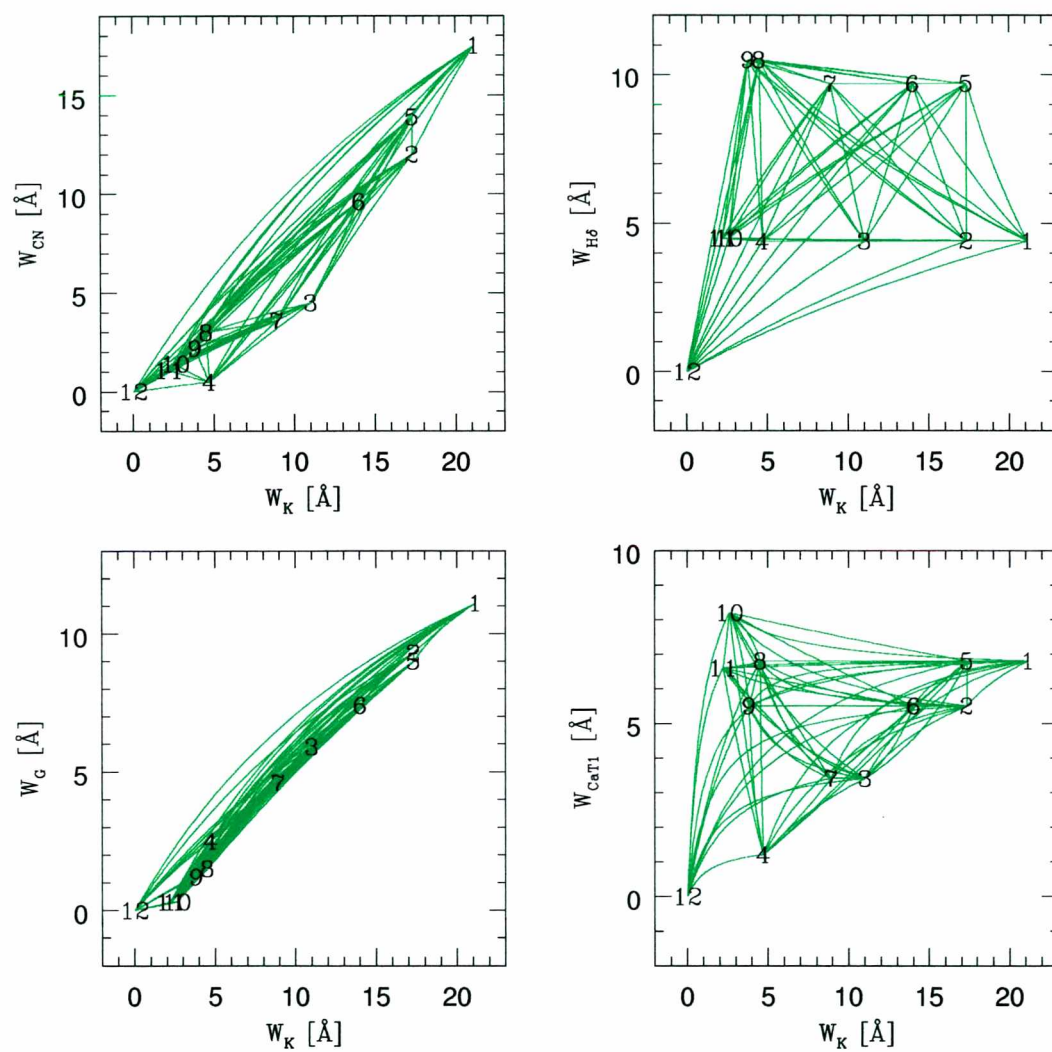


Fig. 6.2: Como a figura anterior, para outras combinações de linhas.

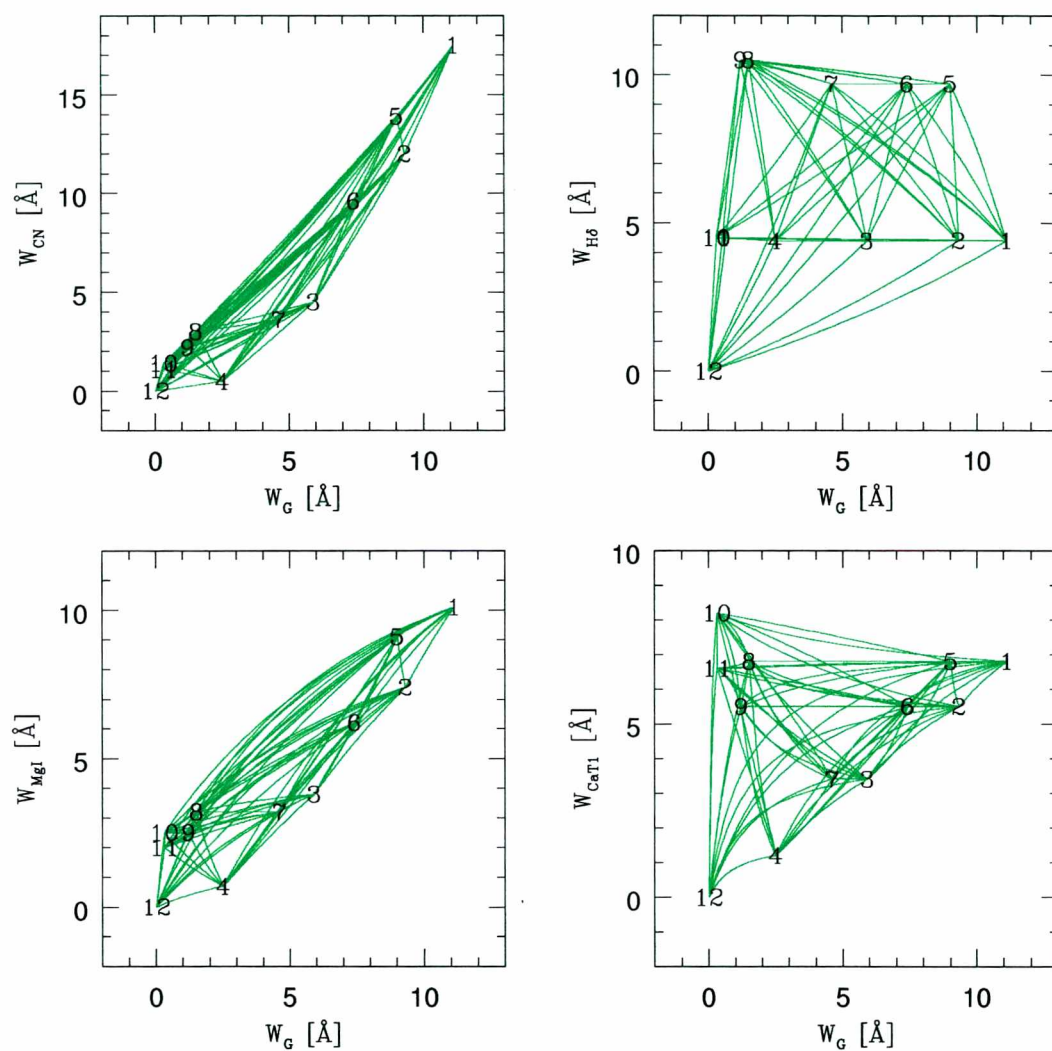


Fig. 6.3: Como a figura anterior, para outras combinações de linhas.

6.2.2 O domínio sintético em diagramas Cor-Cor

Em seu formalismo, Pelat apenas lida com W 's. É possível porém estender o conceito de domínio sintético às cores de maneira trivial. Nas Figuras 6.4 e 6.5 apresentamos algumas projeções do DS em diagramas cor-cor. É importante porém notar que tais diagramas são sensíveis à presença de poeira que avermelha as cores. Diagramas como os das figuras 6.4 e 6.5 podem ser construídos para diferentes valores de A_V , podendo inclusive servir de guia para definir que valores de A_V são necessários para que as cores de uma galáxia sejam sintetizáveis pela base. Na aplicação abaixo nos restringiremos ao uso dos diagramas W - W , pois eles são mais simples e suficientes para nossos propósitos.

6.3 Aplicação às galáxias do atlas de Kennicutt

Na Figura 6.6 apresentamos uma aplicação dos resultados acima às medidas de W 's para 29 galáxias do atlas de Kennicutt (1992). Nas figuras ilustramos o DS nos planos definidos pelas larguras equivalentes das linhas CaII K, CN, banda-G, MgI, CaII $\lambda 8662$, H δ .

O resultado é extremamente interessante: **24 das 29 galáxias não se localizam dentro do DS em todos diagramas!** (exceção de NGC 6181) Em outras palavras, nenhuma dessas galáxias é sintetizável pela base! Isto, claro, em um sentido matemático, algébrico, tal qual Pelat (1997 e 1998) trata o problema.

Tiramos duas conclusões desse breve exercício.

1. A degenerescência algébrica não existe na prática!
2. A base de Schmidt *et al.* (1991) e por extensão a versão original desta base, oriunda do trabalho de Bica (1988), parece ser insuficiente para dar conta de propriedades observadas de galáxias reais! Apenas se admitirmos a presença de erros observacionais pode-se compatibilizar as observações com o DS definido pela base.

O efeito de erros nas W 's e C 's sobre a sintetizabilidade da base ficam claros ao examinarmos as figuras acima. Em vários dos diagramas W - W e C - C o DS é bastante estreito. Uma pequena perturbação devido aos erros observacionais é suficiente para levar uma galáxia de dentro do DS (e portanto sintetizável) para fora dele. O efeito dos erros, portanto, parece ser muito mais fundamental do que o da degenerescência algébrica. Isto dá força ao formalismo apresentado nos capítulos anteriores, onde os erros foram tratados de forma detalhada e consistente.

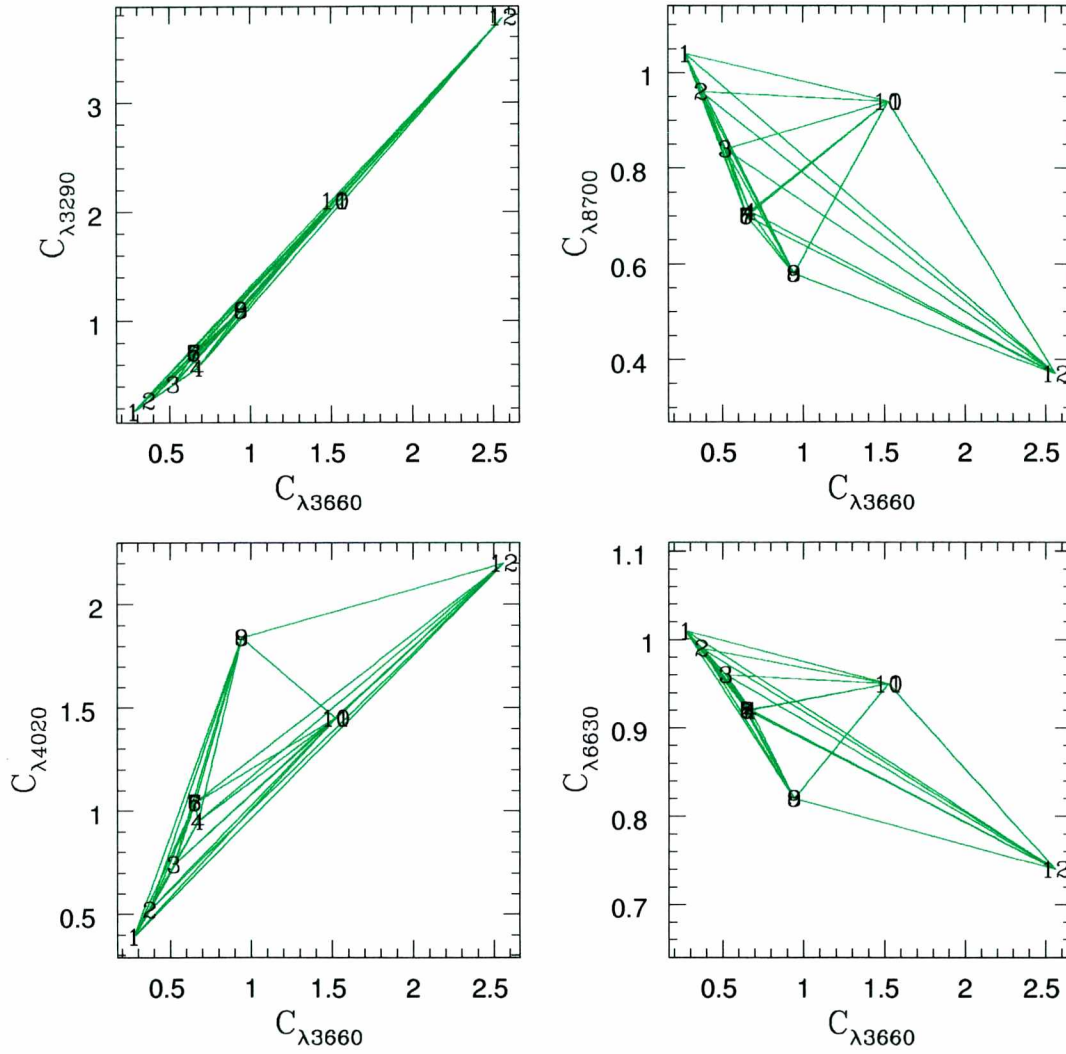


Fig. 6.4: Projeções do domínio sintético no plano definido por fluxos do contínuo, todos normalizados a 5870 Å. As curvas foram calculadas ignorando o efeito de avermelhamento.

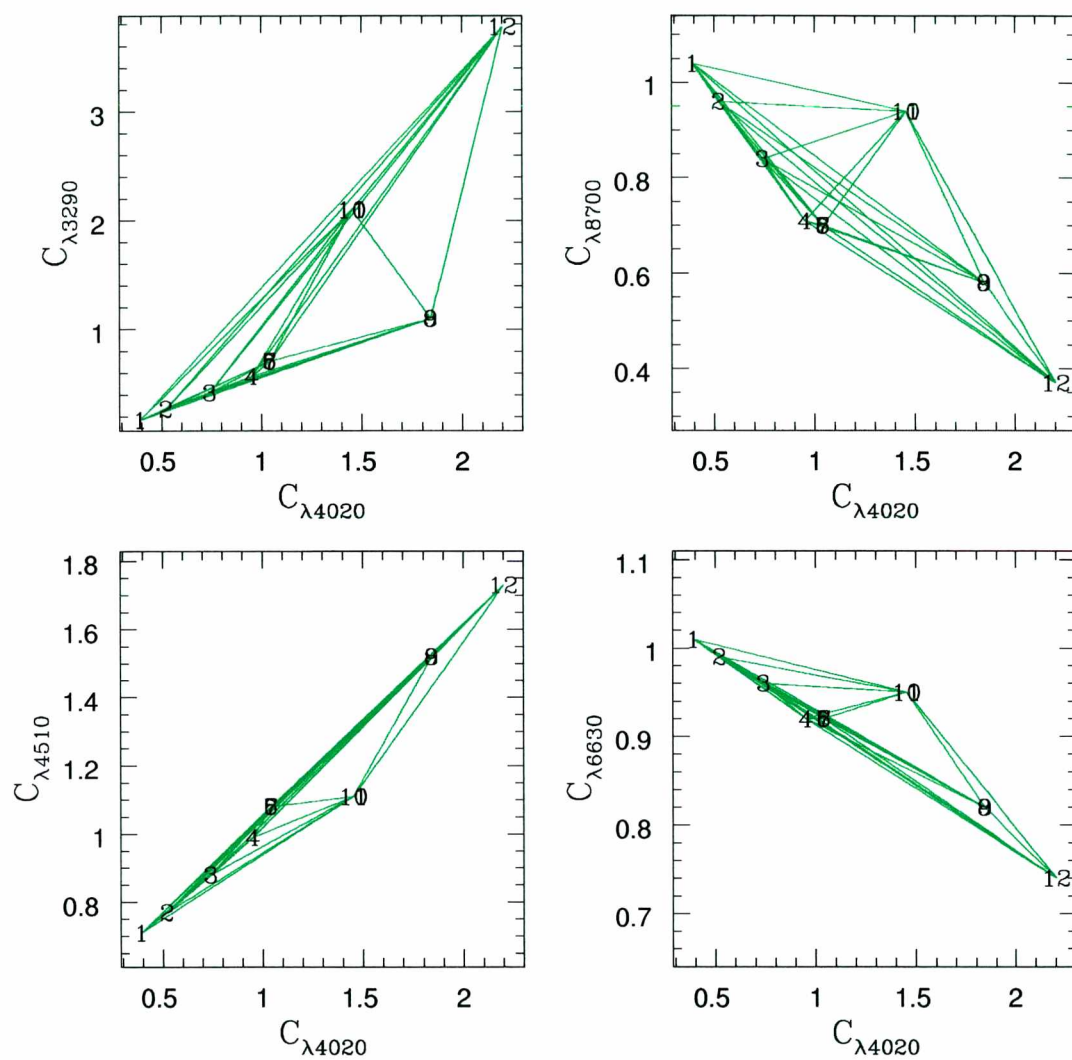


Fig. 6.5: Como a figura anterior, para outras combinações de cores.

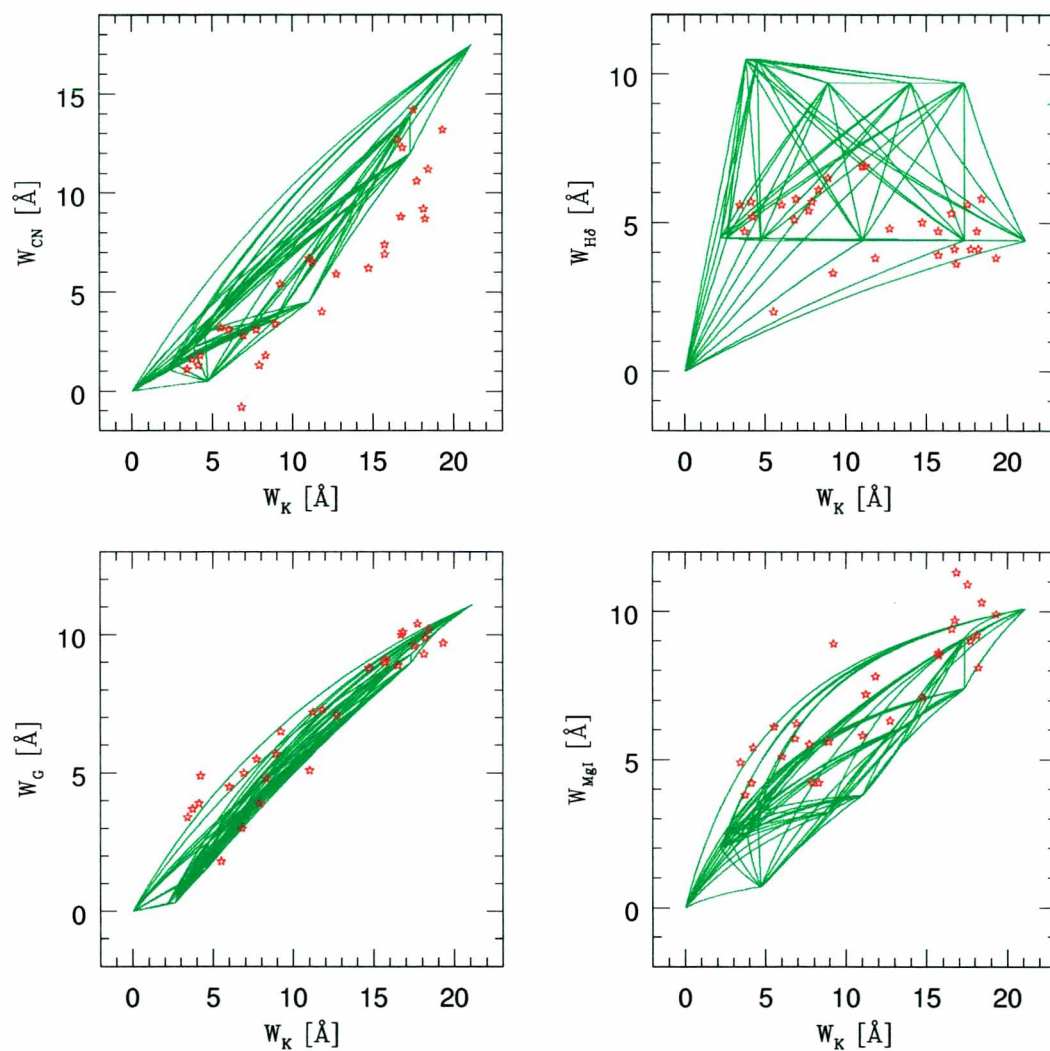


Fig. 6.6: Como a figura anterior, para outras combinações de linhas.

Capítulo 7

Conclusões e perspectivas de continuidade

Revisitamos nesta dissertação o método de Síntese de Populações Estelares Semi-Empírica, com a intenção de contribuir com melhorias em seu formalismo matemático e computacional. Exploramos também sua eficácia em recuperar as misturas de populações presentes nas galáxias. Como uma primeira aplicação do novo método sintetizamos um conjunto de galáxias reais e relacionamos os resultados aos seus aspectos morfológicos. Neste capítulo final vamos apontar nossos principais resultados e conclusões. Além disso, apontamos as possibilidades de futuras aplicações utilizando o método proposto.

7.1 Conclusões finais

Dividimos nossos resultados em cinco partes:

(1) No capítulo 2 propomos uma formulação matemática e estatística ao problema, que tradicionalmente era tratado de maneira menos formal.

(2) Introduzimos o Método de Monte Carlo, baseado no algoritmo de Metropolis, para a amostragem do espaço de parâmetros (capítulo 2) e o testamos no capítulo 3. Mostramos também que o método de amostragem proposto é capaz de mapear de uma maneira mais eficiente o espaço de parâmetros, permitindo que se obtenha curvas de distribuição mais “suaves” e precisas que aquelas obtidas utilizando PCD.

(3) No capítulo 4 aplicamos o algoritmo a galáxias-fictícias construídas a partir da base. Assim, testamos a capacidade da base em recuperar as histórias de formação estelar e enriquecimento químico, que fornecem pistas para traçar a evolução de galáxias. Vimos como os resultados são afetados pelos efeitos dos erros observacionais e pela limitação de dados (conjunto de observáveis). As comparações entre dados de entrada e saída foram feitas utilizando-se as soluções médias (\bar{x} , $\overline{A_V}$), assim como em trabalhos anteriores sobre SPESE. Vimos ainda que a síntese, assim como formulada neste trabalho e utilizando esta base, está sujeita a diversos efeitos:

(3.1) A presença dos erros nos observáveis gera dependências lineares entre os elementos da base. Isto se reflete na redistribuição das probabilidades entre componentes de mesma idade das componentes com grande contribuição para o espectro integrado de um objeto. Surpreendentemente, este efeito é tão significativo, que para espectros de qualidade realística, não se pode inferir nada a respeito das estimativas para o vetor de populações \vec{x} .

(3.2) Reduzir a qualidade dos dados e sintetizar conjuntos com menos informações (apenas larguras equivalentes) a respeito do espectro, produz resultados menos precisos, fazendo com que as curvas de distribuição para os x_i 's se desloquem e se tornem mais largas. Quando se sintetiza as cores juntamente com as larguras equivalentes, surge o avermelhamento como um 13º parâmetro a ser determinado. Sua inclusão entretanto é compensada pelos resultados mais precisos que obtemos para as componentes do vetor de populações \vec{x} . Além disso, A_V não é afetado pelos efeitos sistemáticos apontados em (3.1),

(3.3) Detectamos que é extremamente difícil recuperar o vetor de populações mas que os observáveis são reproduzidos com extrema precisão pela solução média, mesmo para espectros realísticos ($S/N \sim 30$). A origem desta dificuldade não é a degenerescência algébrica, uma vez que o mesmo ocorre quando sintetizamos conjuntos com mais observáveis que parâmetros e para erros observacionais pequenos. Em ambos os casos a verossimilhança se redistribui entre os componentes da base, comprometendo a solução média para os x_i 's. Este efeito tem sua origem no fato de que a própria base possui elementos que podem ser “construídos” como combinações lineares de outros elementos. Ou seja: a base que por definição deveria possuir elementos linearmente independentes, na prática não obedece a esta condição, conforme mostramos na seção 4.1.

(3.4) Agrupamos os resultados em idade e obtivemos resultados muito melhores para os \bar{x} com este procedimento. O mesmo pode ser feito para interpretar outros estudos já publicados utilizando esta base. Entre os grupos de 10^9 e 10^{10} anos, ainda persiste alguma confusão devido ao efeito da degenerescência entre idade e metalicidade conforme discutimos na seção 4.4. Pode-se também agrupar os resultados para as 4 metalicidades compreendidas pela base, mas as interpretações resultantes são comprometidas porque os resultados deste procedimento são menos precisos do que as proporções agrupadas em idade. Isto indica que a base é adequada para estudos evolutivos, mas é extremamente limitada quando se tenta inferir a respeito do enriquecimento químico dos objetos.

(4) No capítulo 5 aplicamos o método a um conjunto de galáxias reais e sintetizamos 25 das 55 galáxias presentes no atlas de Kennicutt (1992). Não interpretamos individualmente os resultados, mas procuramos relacionar os resultados gerais da síntese com a morfologia e o tipo espectral, utilizado na técnica de ACP. Obtivemos

resultados conhecidos a respeito dos objetos, pois verificamos traços de populações jovens em galáxias espirais e fortes contribuições de populações velhas para galáxias elípticas. O mesmo pode ser verificado para o tipo espectral, indicando uma correspondência muito forte entre síntese e ACP. Isto sugere que estas duas técnicas possam ser no futuro “calibradas”. Estes resultados confirmam a plausibilidade do método descrito neste trabalho e indicam possíveis aplicações futuras do método, que serão discutidas na seção 7.2

(5) Finalmente, no capítulo 6, calculamos e representamos graficamente o domínio sintético, recentemente introduzido por Pelat (1998). Através de projeções do domínio sintético em planos de observáveis compreendidos pela base, podemos verificar a sintetizabilidade dos objetos. Em especial, verificamos como as galáxias do atlas de Kennicutt se encontram nestes planos, e verificamos a impossibilidade de sintetizá-las com precisão. O conceito de domínio sintético é, desta forma, bastante útil, pois possibilita verificar a possibilidade ou o grau de confiança dos resultados da síntese, antes mesmo do procedimento.

7.2 Aplicações futuras e perspectivas de continuidade

Trabalhos futuros utilizando este método podem ser feitos em duas frentes: o próprio método e aplicações.

(1) **Melhorias no método:** Os erros observacionais e a convenção de se representar os resultados por soluções médias são responsáveis pelas dificuldades em se recuperar com precisão as populações estelares. Técnicas de minimização, entretanto, podem reduzir estas dificuldades, que reduzem o poder de resolução da base às suas cinco idades. Podemos ainda implementar e utilizar outros Priors, na formulação probabilística que propomos. Por exemplo, a máxima entropia é um destes priores que tem sido utilizado com sucesso em problemas inversos, como recuperação de imagens. A modelagem do espectro por inteiro e a utilização de outros vínculos, assim como a ralação massa-luminosidade (M/L), são outras tentativas a serem feitas.

Apesar das melhorias no método serem possíveis, encontramos que a maior necessidade seria aumentar o poder de resolução em idade e metalicidade da base. Corroboramos portanto com a posição de Schmidt *et al.* (1989 e 1991), que afirmam ser a base a origem da grande maioria das dificuldades apontadas neste trabalho.

(2) **Aplicações:** Podemos imaginar uma grande variedade de empregos para o método. Já neste estágio de desenvolvimento poderíamos aplicá-lo a grandes bancos de dados.

Em particular, a aplicação a certos tipos de fenômenos seria bastante interessante: sintetizar espectros de galáxias em colisão, com o intuito de detectar e datar (pela síntese) eventos de formação estelar abruptos associados ao choque.

A aplicação a núcleos ativos, poderia apontar se a atividade nuclear está associada à formação de estrelas nas vizinhanças do núcleo. De fato, Cid Fernandes *et al.*

(2001b) recentemente fizeram isso (utilizando este método) e encontraram vários indícios de que os dois fenômenos estão relacionados.

Como vimos no capítulo 5 deste trabalho, a concordância entre síntese e ACP é bastante interessante e indica ser possível calibrar os resultados de uma ACP com os da síntese. Isto é interessante porque ACP é automática e pode ser diretamente aplicada a grandes bancos de dados de espectros, tais como a “SLOAN Varredura Digital do céu” (York *et al.* 2000) ou a “Varredura de 2 graus de galáxias a grandes distâncias” (Folkes *et al.* 1999).

Desta forma, vemos que são muitas e imediatas as aplicações e possibilidades desta metodologia. A busca por métodos de análise de espectros e de síntese de populações deve ser alvo constante de futuros trabalhos, pois é imensa a quantidade de dados que existem e maior ainda a quantidade daqueles que estão por vir com os novos observatórios.

Apêndice

A seguir apresentamos o trabalho que submetemos à revista “Monthly Notes of the Royal Astronomical Society (MNRAS)”. O trabalho foi aceito para publicação como artigo regular: Cid Fernandes R., Sodré L., Schmitt H. R., Leão J. R. S., 2001, MNRAS, 325, 60. Este trabalho abrange basicamente os 4 primeiros capítulos desta dissertação.

A probabilistic formulation for Empirical Population Synthesis: Sampling methods and tests

Roberto Cid Fernandes^{1,2,*†}, Laerte Sodr  Jr.^{3‡}, Henrique R. Schmitt^{4§},
Jo  R. S. Le o^{2¶}

¹ *Depart. of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, USA*

² *Depto. de F sica - CFM - Universidade Federal de Santa Catarina, CP 476, Campus Universit rio, Trindade, 88040-900, Florian polis, SC, Brazil*

³ *Depto. de Astronomia, Instituto Astron mico e Geof sico - USP, Av. Miguel Stefano 4200, 04301-904 S o Paulo, Brazil*

⁴ *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*

13 March 2001

ABSTRACT

We revisit the classical problem of synthesizing spectral properties of a galaxy using a base of star clusters, approaching it from a probabilistic perspective. The problem consists of estimating the population vector \mathbf{x} , composed by the contributions of n_* different base elements to the integrated spectrum of a galaxy, and the extinction A_V , given a set of absorption line equivalent widths and continuum colors. The formalism is applied to the base of 12 elements defined by Schmidt et al. (1991) as corresponding to the principal components of the original base employed by Bica (1988), and subsequently used in several studies of the stellar populations of galaxies. The exploration of the 13-D parameter space is carried out with a Markov chain Monte Carlo sampling scheme based on the Metropolis algorithm. This produces a smoother and more efficient mapping of the $P(\mathbf{x}, A_V)$ probability distribution than the traditionally employed uniform-grid sampling.

This new version of Empirical Population Synthesis is used to investigate the ability to recover the detailed history of star formation and chemical evolution using this spectral base. This is studied as a function of (1) the magnitude of the measurement errors and (2) the set of observables used in the synthesis. Extensive simulations with test galaxies are used for this purpose. Emphasis is put on the comparison of input parameters and the *mean* \mathbf{x} and A_V associated with the $P(\mathbf{x}, A_V)$ distribution. It is found that only for extremely low errors ($S/N > 300$ at 5870  ) all 12 base proportions can be accurately recovered, though the observables are recovered very precisely for any S/N . Furthermore, the individual $\bar{\mathbf{x}}$ components are *biased* in the sense that components which carry a large fraction of the light tend to share their contribution preferably among components of same age. Old, metal poor components can also be confused with younger, metal rich components due to the age \times metallicity degeneracy. These compensation effects are linked to noise-induced linear dependences in the base, which very effectively redistribute the likelihood in \mathbf{x} -space. The *age* distribution, however, can be satisfactorily recovered for realistic S/N (~ 30). We also find that synthesizing equivalent widths and colors produces better focused results than those obtained synthesizing only equivalent widths, despite the inclusion of the extinction as an extra parameter.

Key words: galaxies: evolution - galaxies: stellar content - galaxies: statistics

* e-mail: cid@fsc.ufsc.br

† Gemini fellow

‡ e-mail: laerte@iagusp.usp.br

§ e-mail: schmitt@stsci.edu

¶ e-mail: joao@fsc.ufsc.br

1 INTRODUCTION

The stellar populations of galaxies carry a record of their star-forming and chemical histories, from the epoch of formation to the present. Their study thus provide a powerful tool to explore the physics of galaxy formation and evolution. Despite the detailed understanding of stellar evolution and of the properties of single stellar populations such as star clusters achieved over the past century, uncovering the history of galaxies through the study of their integrated light has proven to be a difficult problem (Pickles 1985; Worthey 1994 and references therein).

Two global approaches have been developed to tackle this issue. The first of these is Evolutionary Population Synthesis, which performs *ab initio* calculations of the spectral evolution of galaxies on the basis of stellar evolution theory, stellar spectral libraries, and prescriptions for the initial mass function, star formation rate and chemical evolution (Tinsley 1972; Larson & Tinsley 1974, 1978; Guiderdoni & Rocca-Volmerange 1987; Fioc & Rocca-Volmerange 1997; Charlot & Bruzual 1991; Bruzual & Charlot 1993; Leitherer et al. 1996, 1999 and references therein). The other approach is Empirical Population Synthesis (EPS hereafter), also known as ‘Stellar Population Synthesis with a Database’, which uses observed properties of stars or star clusters as a *base* on which to decompose the population mix in galaxy spectra (Spinrad & Taylor 1971; O’Connell 1976; Faber 1972; Pickles 1985; Bica 1988; Pelat 1997, 1998; Boisson et al. 2000). Of particular interest is the EPS method of synthesizing the spectra of galaxies using a base of star clusters. Perhaps the most appealing property of this method is the fact that the results do not hinge on assumptions about stellar evolution tracks nor initial mass functions, as these are ‘mother nature given’ in the empirically built base spectra. This technique has been pioneered by the work of Bica (1988, hereafter B88), where the stellar content of a large sample of galaxies was synthesized using the spectral base of star clusters described in Bica & Alloin (1986a, 1986b, 1987). Since then, several studies made use of this technique (Schmidt, Bica & Dottori 1989; Bica, Alloin & Schmidt 1990; Jablonka, Alloin & Bica 1990; Schmidt, Bica & Alloin 1990; Jablonka & Alloin 1995; De Mello et al. 1995; Bonatto et al. 1996, 1998, 1999, 2000; Schmitt, Storchi-Bergmann & Cid Fernandes 1999; Kong & Cheng 1999; Raimann et al. 2000).

In its original version (B88), this method made use of a large base of $n_* = 35$ elements, combined in different proportions (x_i ; $i = 1 \dots 35$) to synthesize a number of equivalent widths (W) of conspicuous absorption features. A uniform sweep of the parameter space along paths consistent with simple predefined chemical evolution scenarios was performed, and all combinations which produced W ’s within 10% of the observed values were considered ‘solutions’. The final population vector \mathbf{x} was taken to be the *arithmetic mean* of all such sampled solutions.

One of the main concerns raised by B88 method is that of the *uniqueness* of the solution, as first discussed by Schmidt et al. (1991, hereafter S91). With a 35-D population vector and at most 9 W ’s to be synthesized, one has a highly degenerate algebraic system. As shown by S91, however, the original base was highly redundant, with several linearly dependent elements, and others which were practically so, as deduced by a principal component analysis. S91

then proposed the use of a reduced base of 12 elements. S91 also criticized B88’s uniform sampling of the \mathbf{x} space (a ‘discrete combination procedure’ in their terms) and the use of a ‘mean solution’. Instead, they developed a constrained minimization procedure, which, as shown by their simulations with test galaxies built out of the base, is able to retrieve with good accuracy the correct population vector. Another difference of the S91 method compared to B88 is the possibility to search for solutions along the entire age \times metallicity plane. B88 assumed that the chemical enrichment of the galaxy occurred during its formation (~ 10 Gyr ago), with all stars younger than that having the same Z as the oldest most metal rich component. Although this may be a reasonable approximation for the nuclear region of isolated galaxies, it does not allow for the effects of mergers, inflows or outflows, and is not applicable to high redshift sources, where the spectra usually integrate over the entire galaxy, including regions with different star formation and chemical histories (Jablonka et al. 1990; Jablonka & Alloin 1995).

The next improvement in this technique was the introduction of continuum colors in the synthesis. This was done by Schmitt, Bica & Pastoriza (1996), who, as B88, use a discrete combination procedure search for acceptable solutions, sweeping the parameter space along chemical enrichment paths. This version of EPS has been used in several subsequent studies (Bonatto et al. 1996, 1998, 1999, 2000; Kong & Cheng 1999). Another step was taken by Bonatto, Bica & Alloin (1995), who extended the star cluster base to the space ultraviolet (1200–3200 Å). They too use an arithmetic average to represent the solution of the synthesis.

In two recent papers, Pelat revisited the EPS problem of synthesizing W ’s using an elegant formalism based on convex algebra concepts. In Pelat (1998) he shows how the algebraic solution of the EPS problem can be narrowed down to a sub-space of \mathbf{x} delimited by a set of ‘extreme-solutions’, any convex combination of which yields an exact solution in the underdetermined cases (more parameters than constraints). His results shed new light into the issue of algebraic degeneracy, which has always haunted EPS due to the widespread belief that with more parameters than observables one is bound to fit the data in one way or another. He proposes a quick test of whether the W ’s of a galaxy are synthesizable by computing the region in W -space spanned by all physical combinations of the base elements. We verified that in the case of the 12 elements base used by S91 this region is narrow (Leão 2001), such that small measurement errors are enough to place objects outside this ‘synthetic domain’, hence preventing the existence of mathematically exact solutions. In real applications, one therefore expects the degeneracy of EPS to be more of a statistical nature than algebraic. In these cases, as well as in overdetermined problems, Pelat (1997) demonstrates that the best model is to be found along the boundary of the synthetic domain by minimizing an appropriately defined χ^2 -like distance. Uncertainties in the population vector can also be readily computed, as shown by Moulata & Pelat (2000).

Boisson et al. (2000) recently applied Pelat’s method to a sample of 12 active galaxies from Serote Roos et al. (1998). They synthesize $n_W = 47$ absorption lines with a base of $n_* = 30$ stars from the stellar spectral libraries of Serote Roos, Boisson & Joly (1996) and Silva & Cornell (1992). Their results nicely illustrate the power of this new EPS

technique, which will certainly be extended to larger samples and other classes of galaxies in future studies. The B88 method and its variants, on the other hand, already have a large body of literature associated to it. Nevertheless, this popular and attractive EPS technique, while intuitively acceptable, has neither been adequately tested nor cast onto a sound mathematical formalism which supports its application. Furthermore, it gives no measure of the uncertainties or potential biases on the population vector, thus giving no simple way to prevent overinterpretations of the resulting population of the age \times metallicity plane. These issues are the focus of the present paper.

Our main goals are to:

- (i) Elaborate a mathematically consistent formulation of the EPS problem in the context of probability theory.
- (ii) Develop and test a sampling method to explore the parameter space in EPS problems. In particular, we aim to improve upon the Schmitt et al. (1996) technique of synthesizing both equivalent widths and colors, but within the context of a probabilistic formulation.
- (iii) Use the method to investigate the effects of measurement errors and of the specific set of observables used upon the results of the synthesis with the reduced B88 star cluster base described in S91. In particular, we wish to evaluate whether these factors introduce biases in the population mix inferred from this popular synthesis technique and at what level of detail can its results be trusted.

This paper is organized as follows: Item (i) of the above list is addressed in section 2. In section 3 we deal with point (ii). This is done by investigating the applicability of a Metropolis algorithm to sample the parameters probability distribution in EPS, and testing the method with simulations based on Bica's spectral base. In section 4 we address item (iii) by means of extensive simulations which explore the ability to recover the detailed star-formation and chemical histories of galaxies with this base. Finally, section 5 summarizes our main results.

2 FORMALISM

2.1 Basic equations

EPS studies seek to find combinations of a spectral base which reproduce a given set of measured observables, often taken as the equivalent widths W_j of n_W conspicuous absorption features. The base consists of n_* elements representing well defined simple stellar populations such as star-clusters, with equivalent widths W_{ij}^* and corresponding continuum fluxes over the lines F_{ij}^* ($j = 1, \dots, n_W$, $i = 1, \dots, n_*$) normalized at a reference wavelength. Denoting by x_i the fractional contribution of the i -th base element to the total flux at the reference wavelength, one obtains a system of n_W equations

$$W_j = W_j(\mathbf{x}) = \frac{\sum_{i=1}^{n_*} W_{ij}^* F_{ij}^* x_i}{\sum_{i=1}^{n_*} F_{ij}^* x_i}; \quad j = 1, \dots, n_W \quad (1)$$

for the synthetic W 's (e.g., July 1974). This set of constraints can be augmented by synthesizing n_C observed continuum fluxes C_k ($k = 1, \dots, n_C$), provided allowance is made for reddening, here parametrized by the V-band extinction A_V .

Assuming all n_* populations are equally reddened, one obtains

$$C_k = C_k(\mathbf{x}, A_V) = g_k(A_V) \sum_{i=1}^{n_*} C_{ik}^* x_i; \quad k = 1, \dots, n_C \quad (2)$$

where a distinction is made between C^* and F^* , since the n_C fluxes to be synthesized need not correspond to the wavelengths of the n_W absorption lines. We hereafter refer to the C_k 's as continuum *colors*, as they are in fact ratios of continuum fluxes with respect to the continuum at the normalization wavelength. The function $g_k(A_V)$ reddens the normalized color at wavelength λ_k by A_{λ_k} according to a specified reddening law.

Finally, physical solutions must further satisfy the normalization and positivity constraints:

$$\sum_{i=1}^{n_*} x_i = 1; \quad x_i \geq 0 \text{ for } i = 1, \dots, n_*, \text{ and } A_V \geq 0. \quad (3)$$

The normalization condition effectively reduces one degree of freedom. When modeling colors, however, one has to introduce A_V as a further parameter, so that the number of parameters still is n_* .

2.2 Probabilistic formulation

The data \mathcal{D} to be modeled are thus composed of a set of $n_{obs} = n_W + n_C$ observables. The measurement errors in each observable, collectively denoted by σ , are assumed to be known. Given these, the problem of EPS is to estimate the population vector \mathbf{x} and the extinction A_V that 'best' represents the data according to a well defined probabilistic model. It is equally important to estimate the uncertainties in the parameters, as these prevent over-interpreting the resulting mixture of stellar populations at a level of detail not warranted by the data or by intrinsic limitations of the base.

The probability of a solution (\mathbf{x}, A_V) given the data \mathcal{D} and the errors σ , is given by Bayes theorem:

$$P(\mathbf{x}, A_V | \mathcal{D}, \sigma, \mathcal{H}) = \frac{P(\mathcal{D} | \mathbf{x}, A_V, \mathcal{H}) P(\mathbf{x}, A_V | \sigma, \mathcal{H})}{P(\mathcal{D} | \mathcal{H})}. \quad (4)$$

\mathcal{H} summarizes the set of assumptions on which the inference is to be made. They include: the mapping between parameters and observables is given by eqs. (1) and (2); \mathbf{x} and A_V must satisfy the constraints expressed in eq. (3); the stellar population in the target galaxy is well represented by the base elements; the observational errors are Gaussian; the reddening law is known.

The likelihood $P(\mathcal{D} | \mathbf{x}, A_V, \sigma, \mathcal{H})$ is a measure of how good (or bad) is the fitting to the data for model parameters \mathbf{x} and A_V . Under the hypothesis of Gaussian errors, the probability of the data given the parameters is

$$P(\mathcal{D} | \mathbf{x}, A_V, \sigma, \mathcal{H}) \propto e^{-\mathcal{E}}, \quad (5)$$

with \mathcal{E} defined as half the value of χ^2 :

$$\begin{aligned} \mathcal{E}(\mathbf{x}, A_V) &= \frac{1}{2} \chi^2(\mathbf{x}, A_V) = \frac{1}{2} \sum_{j=1}^{n_W} \left(\frac{W_j^{obs} - W_j(\mathbf{x})}{\sigma(W_j)} \right)^2 \\ &+ \frac{1}{2} \sum_{k=1}^{n_C} \left(\frac{C_k^{obs} - C_k(\mathbf{x}, A_V)}{\sigma(C_k)} \right)^2 \end{aligned} \quad (6)$$

The likelihood, as the total χ^2 , separates into a W and a color related term, of which only the latter depends on A_V .

The normalizing constant $P(\mathcal{D}|\mathcal{H})$ in Bayes theorem (the ‘evidence of \mathcal{H} ’) is irrelevant to the level of inference discussed here, i.e., the estimation of \mathbf{x} and A_V . It can be used, for instance, to compare different spectral bases.

$P(\mathbf{x}, A_V|\mathcal{H})$ is the joint *a priori* probability distribution of \mathbf{x} and A_V , and states what values the model parameters might plausibly take. For instance, physically acceptable population fractions should have priors that are zero in regions of the parameter space where the constraints of positivity and unit sum are not satisfied. A prior that is uniform on the parameters and includes these constraints is a *non-informative prior*, because it does not impose any constraints on the solutions besides those expressed by eq. (3). This is the prior that will be used in this work.

For a non-informative prior the posterior probability $P(\mathbf{x}, A_V|\mathcal{D}, \sigma, \mathcal{H})$ is simply proportional to the likelihood:

$$P(\mathbf{x}, A_V|\mathcal{D}, \sigma, \mathcal{H}) \propto e^{-\mathcal{E}(\mathbf{x}, A_V)} \quad (7)$$

This expression contains the full solution of the EPS problem, as embedded in it is not only the most probable model parameters but also their full probability distributions. Furthermore, projected posterior distributions for any of the n_* parameters can be obtained by *marginalizing* eq. (7) with respect to all other parameters. For instance, the probability density of proportion x_i of the i -th base element is

$$P(x_i|\mathcal{D}, \sigma, \mathcal{H}) = \int \dots \int P(\mathbf{x}, A_V|\mathcal{D}, \sigma, \mathcal{H}) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_{n_*} dA_V \quad (8)$$

and equivalently for A_V . Similarly, one can construct joint posteriors for, say, the total proportion of all base components of same age or Z , or for the mean age and Z of the stellar population. This would of course provide a coarser description of a galaxy’s stellar content than the individual posteriors $P(x_i|\mathcal{D}, \sigma, \mathcal{H})$, but that may be all that is possible under some circumstances, such as when only a reduced set of observables is available or when observational errors are large (section 4).

It is easy to see how the B88 and Schmitt et al. (1996) synthesis techniques fit into this general probabilistic formulation. By synthesizing the observables within $\sim 10\%$ ‘error boxes’, these authors implicitly assumed a box-car likelihood function $P(\mathcal{D}|\mathbf{x}, A_V, \sigma, \mathcal{H})$, whereas by performing arithmetic means over all accepted (\mathbf{x}, A_V) combinations in a uniform grid they are implicitly sampling the corresponding posterior probability distributions. Also, the *a priori* constraints on the occupation of the age $\times Z$ plane imposed by these authors (but relaxed in subsequent works) simply reflect their use of an informative prior. This gives some formal justification for their heuristically designed EPS method. It is therefore reasonable to expect that an implementation of EPS based on the formalism presented here should give results roughly compatible with those previously obtained. An uniform sweep of the (\mathbf{x}, A_V) space is however a very inefficient sampler for (8), so some improvement is needed there. This is discussed next.

3 SAMPLING METHOD AND TESTS

Despite its formal merits, at the computational level our probabilistic approach faces the same basic difficulty as previous EPS codes, namely, the high dimensionality of the problem. For spectral bases with astrophysically interesting resolution in age and metallicity the number of elements n_* quickly becomes large enough to render the exploration of the parameter space a non-trivial task. Before discussing methods to sample the (\mathbf{x}, A_V) space, we present the spectral base used in this work (§3.1) and describe how we deal with the uncertainties in the observables (§3.2).

3.1 The spectral base and observables

The base used for the synthesis here is that of S91. It contains $n_* = 12$ population groups, spanning five age bins—10 Gyr (which actually represent globular cluster-like populations), 1 Gyr, 100 Myr, 10 Myr and HII (corresponding to current star formation, and represented by a pure $F_\lambda \propto \lambda^{-2}$ continuum based on the spectrum of 30 Dor)—and four metallicities—0.01, 0.1, 1 and 4 Z_\odot (Table 1). The observables in this base comprise the $n_W = 9$ equivalent widths of the absorption lines CaII K $\lambda 3933$, CN $\lambda 4200$, G band $\lambda 4301$, MgI $\lambda 5175$, CaII $\lambda 8543$, CaII $\lambda 8662$, H δ , H γ and H β , as well as $n_C = 7$ continuum fluxes at selected pivot wavelengths: 3290, 3660, 4020, 4510, 6630, 7520 and 8700 Å, all normalized to 5870 Å. According to Bica & Alloin (1986a, 1987) and Bica, Alloin & Schmitt (1994), the equivalent width windows and continuum points were defined based on very high signal to noise spectra of galaxies, with the express goal of using them to synthesize the stellar population of galaxies. These spectra were obtained by creating two average spectra, of blue and red galaxies, where the stellar population features can be clearly traced. The base values for these quantities have been previously published in Bica & Alloin (1986b, 1987) and Schmitt et al. (1996), and some were measured from data in Bica et al. (1994). These are recompiled in Table 2, as they bear a direct impact on the analysis below. There are thus a maximum of $9 + 7 = 16$ observables to model with $n_* = 12$ parameters: $n_* - 1$ population fractions plus A_V . However, in several of the tests presented below we shall make use of a reduced subset of observables, as observational data sets seldom cover the whole spectral range spanned by this base.

According to Bica & Alloin (1986a, 1987) and Bica, Alloin & Schmitt (1994), since the ultimate goal of the continuum points and W ’s is to synthesize the stellar population of galaxies, the W windows and continuum points were defined based on very high signal to noise spectra of galaxies. These spectra were obtained by creating two average spectra, of blue and red galaxies, where the stellar population features can be identified.

The reddening of the colors is modeled with the extinction law described in Cardelli, Clayton & Mathis (1989, with $R_V = 3.1$).

3.2 Errors in the observables

The measurement errors in the observables play a key role in determining the structure of the likelihood function. Whereas such errors are available when analyzing observed

Base Elements Used					
HII	10 Myr	100 Myr	1 Gyr	10 Gyr	$\log(Z/Z_{\odot})$
	10	8	5	1	0.6
12	11	9	6	2	0.0
			7	3	-1.0
				4	-2.0

Table 1. Ages, metallicities and numbering convention for the star clusters in the base.

Equivalent Widths (Å)									
#	K	CN	G	MgI	CaT ₁	CaT ₂	Hδ	Hγ	Hβ
1	21.1	17.5	11.1	10.1	6.8	6.0	4.4	4.9	3.5
2	17.3	12.0	9.3	7.4	5.5	5.0	4.4	4.9	3.5
3	11.0	4.5	5.9	3.8	3.4	3.3	4.4	4.9	3.5
4	4.7	0.5	2.5	0.7	1.2	1.6	4.4	4.9	3.5
5	17.3	13.9	9.0	9.1	6.8	6.0	9.7	7.7	7.5
6	14.0	9.6	7.4	6.2	5.5	5.0	9.7	7.7	7.5
7	8.9	3.6	4.6	3.2	3.4	3.3	9.7	7.7	7.5
8	4.5	3.0	1.5	3.2	6.8	6.0	10.5	9.9	7.9
9	3.8	2.2	1.2	2.5	5.5	5.0	10.5	9.9	7.9
10	2.6	1.4	0.3	2.5	8.2	6.9	4.5	3.5	3.9
11	2.2	1.1	0.3	2.0	6.6	5.8	4.5	3.5	3.9
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Continuum over the lines									
1	0.34	0.48	0.55	0.87	1.03	1.04	0.44	0.58	0.79
2	0.48	0.59	0.65	0.90	0.96	0.96	0.56	0.67	0.83
3	0.72	0.78	0.81	0.95	0.84	0.84	0.76	0.82	0.91
4	0.94	0.96	0.97	1.01	0.72	0.71	0.96	0.97	1.00
5	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
6	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
7	1.03	1.05	1.06	1.04	0.70	0.70	1.05	1.06	1.06
8	1.89	1.75	1.68	1.23	0.59	0.58	1.79	1.65	1.38
9	1.89	1.75	1.68	1.23	0.59	0.58	1.79	1.65	1.38
10	1.55	1.34	1.27	1.06	0.95	0.94	1.40	1.25	1.08
11	1.55	1.34	1.27	1.06	0.95	0.94	1.40	1.25	1.08
12	2.34	2.03	1.93	1.37	0.40	0.38	2.11	1.90	1.56

Continuum Colors								
#	3290	3600	4020	4510	6630	7520	8700	
1	0.17	0.27	0.39	0.71	1.01	1.01	1.04	
2	0.27	0.37	0.52	0.77	0.99	0.96	0.96	
3	0.42	0.52	0.74	0.88	0.96	0.89	0.84	
4	0.57	0.67	0.95	0.99	0.92	0.81	0.71	
5	0.71	0.65	1.04	1.08	0.92	0.81	0.70	
6	0.71	0.65	1.04	1.08	0.92	0.81	0.70	
7	0.71	0.65	1.04	1.08	0.92	0.81	0.70	
8	1.10	0.94	1.84	1.52	0.82	0.68	0.58	
9	1.10	0.94	1.84	1.52	0.82	0.68	0.58	
10	2.10	1.52	1.45	1.11	0.95	0.99	0.94	
11	2.10	1.52	1.45	1.11	0.95	0.99	0.94	
12	3.78	2.56	2.20	1.73	0.74	0.56	0.37	

Table 2. Observables in the base, which make the matrices W_{ij}^* , F_{ij}^* and C_{ij}^* . All continuum fluxes are normalized to the continuum at 5870 Å.

galaxy spectra, they have to be postulated in the synthesis of test galaxies performed below. An alternative and potentially more attractive approach to deal with the errors is to marginalize eq. (8) over σ . This yields more conservative estimates for the parameters, insofar as they are independent

of the actual observational errors. In this paper we follow the traditional approach of treating σ as a relevant constraint in the synthesis process.

In order to insure a realistic correspondence between the values of the errors and the actual quality of the data

have adopted a prescription for the σ 's based on our experience in dealing with the measurement of W 's and s in galaxy spectra (Cid Fernandes, Storchi-Bergmann & Schmitt 1998). Two sources of uncertainty affect the measurements of W 's: (1) the noise within the line window, and (2) the uncertainty in the positioning of the pseudo-continuum fluxes C_k at the pivot wavelengths, which affect λ_j because they define the continuum over the line. The effect of the first of these sources upon $\sigma(W_j)$ is straightforwardly obtained through standard propagation of errors by specifying the noise level at λ_j , the spectral resolution $\delta\lambda$, and the size Δ_j of the pre-defined window over which W_j is measured. In Bica & Alloin (1986a) system the pseudo-continuum is defined interactively by visual inspection of the spectrum around the pivot λ 's, which makes the estimation of $\sigma(C_k)$ non straight-forward. In Cid Fernandes et al. (1998) we found these errors could be roughly scaled from the signal-to-noise ratio measured in nearby 'line-free' regions, such that the S/N in the pseudo continuum is $(S/N)_{PC} \sim \rho_{PC}(S/N)_\lambda$, with $\rho_{PC} \sim 2-3$. This leads to the following expression for the $\sigma(W_j)$'s:

$$^2(W_j) = \Delta_j \delta_\lambda (S/N)_\lambda^{-2} + (\Delta_j - W_j)^2 \rho_{PC}^{-2} (S/N)_\lambda^{-2} \quad (9)$$

where the two terms correspond to the two sources of uncertainty discussed above. For simplicity, we further assume that the noise in the observed spectrum is constant for all λ 's, which results in $(S/N)_\lambda = C_\lambda (S/N)_{5870}$. The errors in the C_k 's thus become also constant:

$$\sigma(C_k) = \rho_{PC}^{-1} (S/N)_{5870}^{-1} \quad (10)$$

Besides yielding realistic values for the σ 's, this recipe has the advantage of quantifying the quality of the data in terms of a single quantity: The signal-to-noise ratio at the normalization wavelength 5870 Å. Of secondary importance are the spectral resolution and ρ_{PC} , which in the simulations below are kept fixed at 5 Å and 3 respectively.

In order to provide a quantitative notion of how a value of S/N translates into $\sigma(W_j)$'s, we have computed these quantities for the observables corresponding to spectral groups S1 and S7 of B88, typical of the nuclei of early (red) and late (blue) type spiral galaxies respectively. For S1 our error recipe yields $\sigma(W_{CaIIK}) = 36 \text{ Å} / (S/N)_{5870}$ and $\sigma(W_{MgI}) = 19 \text{ Å} / (S/N)_{5870}$, so errors of $\sim 1 \text{ Å}$ or less are achieved with $(S/N)_{5870} \geq 30$. For S7 these lines have $\sigma(W) \sim 15 \text{ Å} / (S/N)_{5870}$.

3.3 Exploration of the parameter space: Sampling methods

The exploration of the (x, A_V) space can be approached from two alternative perspectives: (1) A *minimization* problem, employing methods to search for the global χ^2 minimum (maximum likelihood). Minimization techniques were explored in S91 and Pelat (1997) for the synthesis of W 's only. (2) A *sampling* problem, where one seeks to map out the posterior probability distribution (eq. 7).

The methods discussed in this paper are *sampling* methods, which do not explicitly search for a minimum. In the limit of small errors, however, one expects the posterior maps to peak at the most likely model. For large errors, on the other hand, the best model sampled may be well off the

true global minimum, but in this case the probability distribution is so broad that the very meaning of 'best model' is questionable. In such cases it is arguably more important to estimate plausible ranges for the parameters than to carry out a refined search for the most likely values.

Our main motivation to explore the sampling approach is that most applications of EPS with Bica's base to date focused on the estimation of *mean* parameters, whose meaning cannot be directly ascertained without estimating their uncertainty and/or biases. A critical re-evaluation of this method is thus crucial to establish to which level of detail the results of this popular technique can be trusted.

3.3.1 Uniform sampling

In order to compute the individual posterior probabilities for each parameter, we first note that eq. (8) for the x_i posterior is simply the *mean* probability $P(x, A_V | \mathcal{D}, \sigma, \mathcal{H})$ over the space spanned by all possible values of $x_{j \neq i}$ and A_V for a fixed x_i . The simplest method to compute such probabilities is to divide the (x, A_V) space into a *uniform* grid with Δx steps for the population fractions and ΔA_V for the reddening parameter. One then approximates eq. (8) by the finite sum

$$P(x_i | \mathcal{D}, \sigma, \mathcal{H}) \simeq \frac{\sum_s e^{-\mathcal{E}(x_s, A_{V,s})} \delta(x_i - x_{i,s})}{\sum_s e^{-\mathcal{E}(x_s, A_{V,s})}} \quad (11)$$

where s denotes a point (a 'state') in the 13-D grid and the $\delta(x_i - x_{i,s})$ term retains only those grid points where population i contributes with fraction x_i of the light.

This discrete sweeping of the parameter space is easy to implement, and was in fact the recipe followed by most works with Bica's base to date (see section 1). A serious drawback of uniform sampling is its computational price. The number of grid cells is an extremely steep function of the resolution Δx . A coarse $\Delta x = 10\%$ sampling yields 3.5×10^5 x -points, increasing to $\sim 8 \times 10^7$ for $\Delta x = 5\%$ and more than 4×10^{11} for a 2% resolution! To obtain the grid size one has to further multiply these numbers by the number of points in the A_V grid, whose limits have to be pre-defined. Schmitt et al. (1999), for instance, performed a synthesis study with $\Delta x = 5\%$ and $\Delta A_V = 0.06$ for A_V between 0 and 1.5, which yields 2×10^9 grid points, perhaps the finest grid ever used in EPS with this 12 elements base. One obviously always aim for the best resolution possible, but *a priori* estimates of what resolution is necessary are not straight-forward. A 'good' resolution should produce probability distributions that are smooth over scales of Δx and ΔA_V . The width of these distributions is of course a function of the observational errors and of the set of observables being synthesized.

3.3.2 Metropolis sampling

A further drawback of uniform sampling is that the algorithm spends most of the time in regions of negligible probability, which contribute little to the integral in equation (8), or its numerical version (eq. 11). A more efficient sweep of the parameter space would be to use a sampling scheme which traces the full probability distribution (eq. 7). One

uch ‘importance sampling’ scheme is the Metropolis algorithm (Metropolis et al. 1953), which samples preferentially regions where the posterior is large. This is the parameter-space exploration method adopted in this work.

Our implementation of the Metropolis sampler was as follows. Starting from an initial arbitrary point, at each iteration s we pick one of the 13 variables at random and change it by a uniform deviate ranging from $-\epsilon$ to $+\epsilon$, producing a new state $s + 1$. If the variable is one of the 12 x_i ’s, the whole set is renormalized to unit sum. Moves towards unphysical values ($x_i < 0$ or > 1 , $A_V < 0$) were truncated. Downhill moves (i.e., towards smaller χ^2) are always accepted, whilst changes to less likely states are accepted with probability $e^{-(\mathcal{E}_{s+1}-\mathcal{E}_s)}$, thus avoiding trapping onto local minima. This scheme, widely employed in statistical mechanics (Press et al. 1992; MacKay 2001 and references therein), produces a distribution which tends to the correct one as the number of samples N_s increase. There is, however, no universal prescription to choose optimal values for ϵ or N_s , so some experimentation is needed.

3.4 Performance tests

In order to test our Metropolis EPS code we performed a series of simulations for artificial galaxies generated out of the base. Two sets of test galaxies were used. The first is composed of the 26 galaxies used by S91 (see their Table 5) to test their minimization algorithm. As they did not synthesize colors, we have used a fixed value of $A_V = 2/3$ to redden all their galaxies. The second set is composed of 100 test-galaxies whose population vectors and extinctions were generated randomly using a scheme which insures that in many cases a few components dominate the light. All simulations presented in this paper were performed sampling $N_s = 10^8$ states, with the ‘visitation parameter’ ϵ set to 0.005 for both the x_i ’s and A_V . Experiments were also performed changing these quantities, and found to give nearly identical results. We note that less samples can be used if one is only interested in the mean (\mathbf{x}, A_V) , as this naturally converges much faster than its full probability distribution. Simulations were performed for several values of $(S/N)_{5870}$ and different sets of observables. Here we restrain the discussion to the illustration of the sampling method.

In Fig. 1 we plot the individual projected posterior probability distributions for each of the 12 x_i ’s plus A_V for two of S91’s test galaxies composed of widely different stellar populations. The posteriors were computed synthesizing all 9 lines and 7 colors in the base (‘Set A’, as defined below) and using $(S/N)_{5870} = 1000$ to define the errors in the observables. This is clearly an unrealistic, highly idealized situation, but it serves as a test-bed for the method in the limit of perfect data. The simulations were started from the $x_i = 1/12$ and $A_V = 0.5$ point. The plot shows that the posterior probabilities are sharp functions peaking at the expected input value, marked by upper arrows, demonstrating that the Metropolis sampler converges to the most likely region. One notes, however, that the strong components in the old age bin (such as x_1 , x_2 and x_3 in the ‘old galaxy’ example, drawn as solid lines) have broader individual posteriors. This effect is further discussed in the next section.

In the opposite limit of large errors, $\sigma \rightarrow \infty$ (the ‘infinite temperature’ regime, as it would be called in statistical

mechanics) the likelihood loses its discriminating power and becomes \sim constant over the whole parameter-space. Eq. (8) then tends to $P(x_i) \propto (1 - x_i)^{10}$ for a $n_* = 12$ base. In this regime the data does not constrain the model at all, and this expression for $P(x_i)$ simply reflects the volume of the \mathbf{x} -space spanned by a given value of x_i . Not surprisingly, this distribution peaks at $x_i = 0$, and has a mean value $\bar{x}_i = 1/12$. We have verified that the Metropolis scheme recovers this distribution satisfactorily as $S/N \rightarrow 0$, which serves as a further test of the adequacy of this sampling method. The skewing of several of the individual posteriors plotted in Figs. 3 and 8 towards small x_i is partly due to this effect, which drags the components to an intrinsically larger region of \mathbf{x} -space.

The evolution of the Metropolis walk through the 13-D (\mathbf{x}, A_V) space is illustrated in Fig. 2. For clarity, only 6 components of \mathbf{x} and just one state for each 10^4 steps are shown. Note that projected $P(x_i)$ posteriors such as those in Fig. 1 are essentially histograms of the x_i states sampled in walks as that in Fig. 2, but weighted by their $\exp -\chi(\mathbf{x})^2/2$ likelihood. The input values of the components shown are indicated at the right edge of the plot for comparison. This particular example corresponds to a simulation with $S/N = 60$ and with all 16 observables synthesized. One sees relatively large excursions of the parameters from their true values. Furthermore, anti-correlations between some of the components (x_1 and x_2 , x_8 and x_9) are clearly visible. This ‘mirror effect’ becomes much more pronounced as S/N increases, whereas for low S/N the broadening of the likelihood function washes out such correlations. This is a consequence of redundancies between several of the base components, as further discussed below. The top panels in Fig. 2 show the evolution of the χ^2 and how two of the synthesized observables, $W(\text{CaIIK})$ and C_{4020} , oscillate very much within their respective ± 1 sigma error ranges. One therefore anticipates an excellent fit to the data, and indeed the χ^2 obtained for the *mean* (\mathbf{x}, A_V) ‘solution’ in this example is just 0.8 (see also Fig. 7).

Overall, we conclude that the Metropolis algorithm is an efficient sampler of the parameter space for EPS problems. It provides a much more continuous mapping of \mathbf{x} and A_V than would be possible with a uniform grid. Furthermore, it is substantially more efficient computationally. A uniform grid with 10^8 points would only sample each x_i at $\sim 7\%$ steps and 10 values for A_V . By concentrating on the relevant regions of the parameter space, we achieve a much better resolution, as can be judged by the smoothness of the posterior distributions on scales of 1–2%.

One possible variation of this scheme is to combine the uniform grid and Metropolis routines, starting a Markov chain from each point on a uniform grid. Experiments with this alternative scheme were performed, and found to produce essentially the same results. This is to be expected, since we verified that the starting point has little effect upon the resulting posterior distributions. In fact, we performed a series of simulations using the *input parameters as the starting point* and obtained identical results. Though further improvements are possible, the simple Metropolis sampler already provides a substantial improvement over previous works with this base.

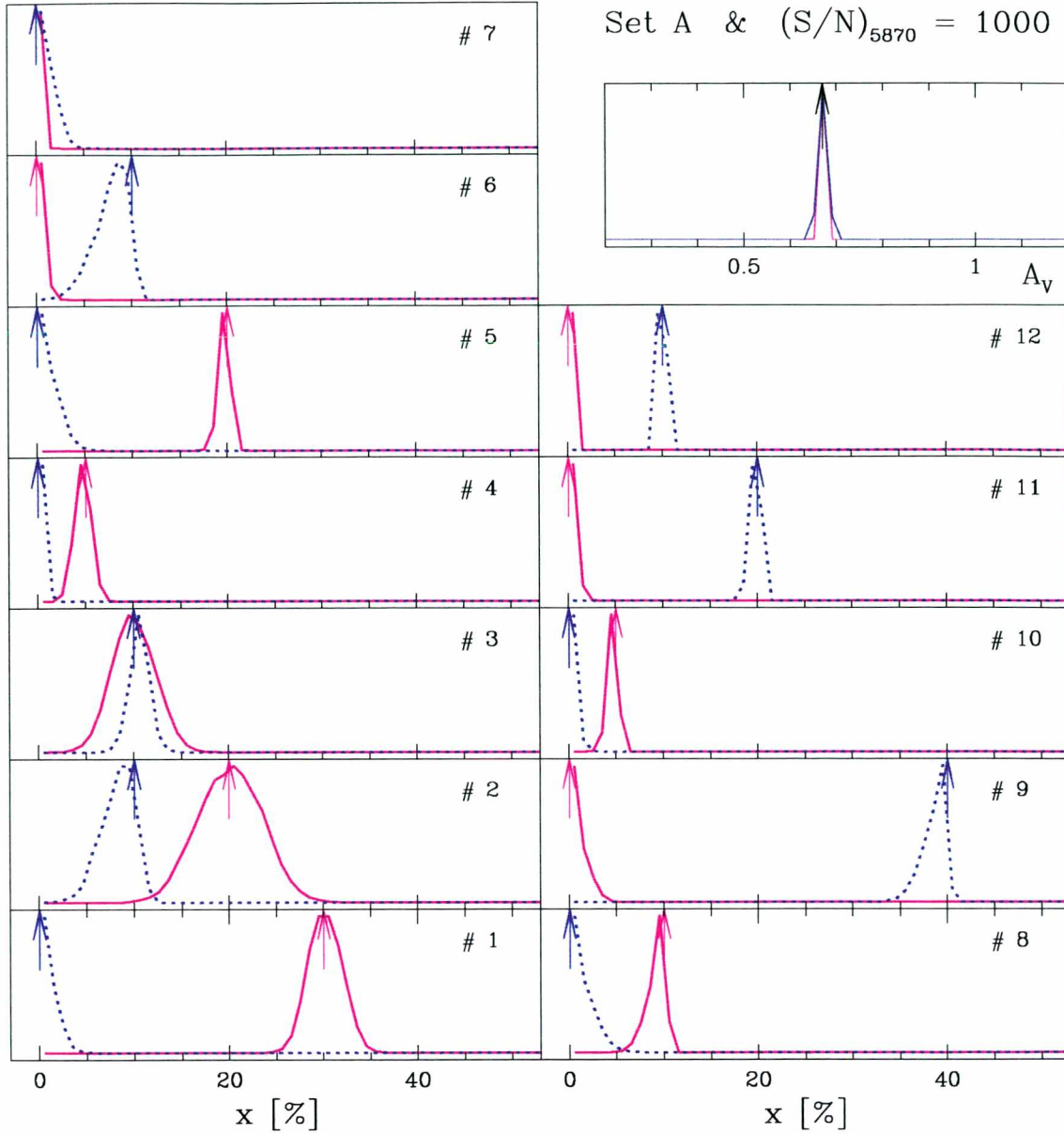


Figure 1. Individual posterior probability distributions for each of the components of the population vector \mathbf{x} and the extinction, computed for an old population dominated test galaxy (solid lines) and a ‘younger’ galaxy (dotted lines). The top arrows indicate the input values of the corresponding parameter in each panel. The numbering of the \mathbf{x} components follows the convention in Table 1. All 16 observables in Bica’s base (9 absorption lines + 7 continuum points) were synthesized in this example and a S/N of 1000 was adopted to illustrate the convergence of the Metropolis sampler in the limit of (almost) perfect data.

4 TESTS OF THE BASE

Having developed a probabilistic formulation of EPS and a new sampling method, we now present a series of numerical experiments designed to evaluate the actual ability to recover the stellar population mix of galaxies using Bica’s base. The tests were performed with the same series of fictional galaxies described in section 3.4. Simulations were

performed for $(S/N)_{5870} = 10, 30, 60, 100, 300$ and 1000, and three different sets of observables:

- Set A: All 16 observables in the base.
- Set B: Only the 9 absorption line equivalent widths.
- Set C: The W ’s of CaII K, CN, G-band and MgI+MgH plus the continuum fluxes at 3660, 4020, 4510 and 6630, all normalized to 5870 Å.

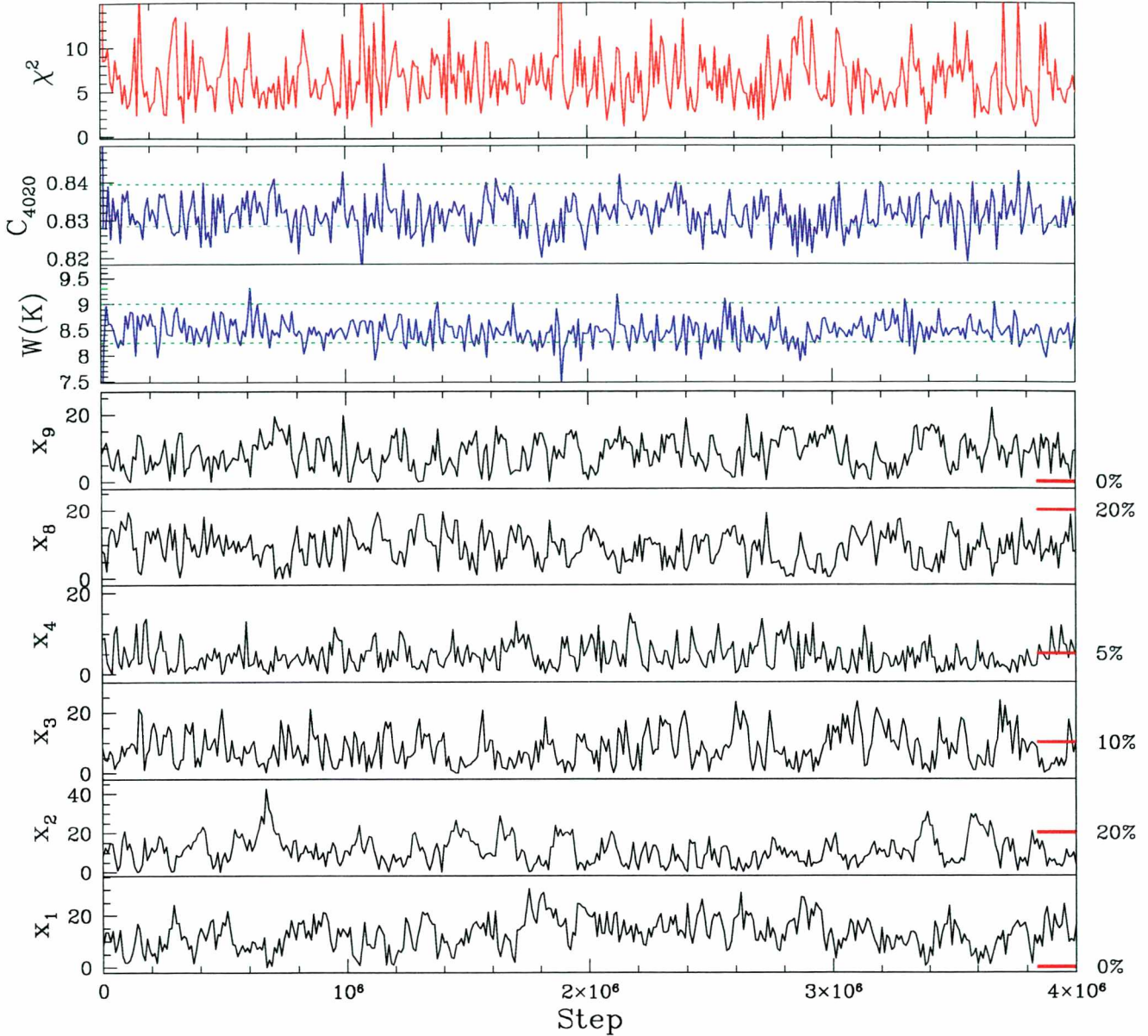


Figure 2. Illustration of the ‘step by step’ Metropolis walk through the parameter space. The six bottom panels illustrate the evolution of selected components of the population vector (in %), whose input x_i values are indicated at the right edge of each plot. The values of the equivalent width of CaIIK and of the color C_{4020} are plotted for each step in the third and second panels from the top. The dotted horizontal lines in these plots mark the $\pm 1\sigma$ range around the observed value. The top panel shows the χ^2 for each sampled state. For clarity, only the first few million steps are shown, and only 1 for every 10000 sampled states is actually plotted. All 16 observables in the base and $S/N = 60$ were used in this example.

Set A is the ideal, as it uses all information in the base. Set B is used as a test of how much is actually gained by synthesizing colors as well as W ’s, while Set C is composed of the observables used by Cid Fernandes et al. (1998) and Schmitt et al. (1999) to characterize the stellar content of active galaxies and their hosts through long-slit spectroscopy. These data will be used to perform a spatially resolved EPS study in a future communication.

4.1 Effects of the errors in the observables

We first investigate the effects of the measurement uncertainties in the observables upon the results of the synthesis. Qualitatively, one expects the degradation of the data quality to *broaden* the probability distributions. Furthermore, a *shift* and *skewing* of the $P(x_i)$ ’s of intrinsically large components towards small x_i ’s is also expected for low S/N due to

intrinsically larger number of states in this region of \mathbf{x} (the ‘ragging effect’ discussed in §3.4). This is exactly what is served in Fig. 3, where we overplot the individual posteriors for $S/N = 1000$, 100 and 10 for one particular test galaxy of S91, synthesized with Set A observables. As in Fig. 1, the true parameters are indicated by the top arrows. The broadening of posteriors in this case can be safely attributed to the errors alone, since, as demonstrated by Pelat (1998) most of the test galaxies in S91 (including the one in Fig. 3) have a unique algebraic solution in a W -only EPS problem, despite the fact that the number of degrees of freedom (11 if we do not model the colors) falls short of the number of observables (9 W ’s).

Despite the large number of constraints in Set A, one sees that the individual posterior distributions are substantially broad even for a $S/N = 100$ spectrum. The average standard deviation of the x_i ’s for this combination of S/N and observables is 3% over all simulations, but can reach more than 10% in individual components. Furthermore, the dominant components, 2, 6 and 9 in this example, are all badly affected by the errors. Fig. 3 shows that decreasing data quality progressively shifts $P(x_2)$ towards smaller values, whereas the opposite happens with x_1 . This *compensation* is also seen between components in the 1 Gyr (x_5 , x_6 and x_7), 100 Myr (x_8 and x_9), and 10 Myr (x_{10} and x_{11}) age bins and is the same ‘mirror’ effect observed before in Fig. 2. A compelling visualization of these compensations is given in Fig. 4, where we plot the individual Metropolis states for the test galaxy employed in Fig. 3 in a $x_i \times x_j$ matrix. The S/N of 300 in this plot was chosen for clarity purposes; the same structure is present for other values, with the scatter around the strongest correlations increasing steadily as S/N decreases, to the point that at $S/N = 10$ the anti-correlations between components 1 and 2, or 8 and 9 become clouds similar to $x_1 \times x_5$ in Fig. 4. All anticorrelations identified in Fig. 3 ($x_1 \times x_2$, $x_6 \times x_5$, etc.) are clearly represented in this graphical version of the correlation matrix.

These effects are rooted in the internal structure of the base. By definition, the base elements must be linearly independent, and the S91 base complies with this algebraic condition. Yet, some of its elements are *practically* linearly dependent, in the sense that combinations of other elements can recover them to a high degree of accuracy. As the noise increases, the residuals between representing an element by its exact W ’s and C ’s and a combination including other elements with more extreme values for the observables become statistically insignificant, explaining the compensations detected above. To demonstrate this we synthesized the base elements themselves with our code. Results for Set A observables and three different values of S/N are shown in Fig. 5. Each panel corresponds to one of these 12 extreme test galaxies (labeled B1...B12), with the *mean* synthetic values of each of the components plotted along the vertical axis. The empty, shaded and filled histograms correspond to $S/N = 100$, 30 and 10 respectively. Some of the components, most notably numbers 2, 3 and particularly 6, are synthesized with large contributions of others even for $S/N = 100$. For the $x_6 = 100\%$ model (top left panel of Fig. 5), for instance, we find $\bar{x}_6 = 60\%$ for this high S/N , with 35 of the remaining 40% redistributed among components 5 and 7. Yet, the 16 observables are very well reproduced by this

combination, with a total χ^2 of just 3.8. This spreading of the light fractions becomes more pronounced for lower S/N ’s, to the point that at $S/N = 10$, components like 8 and 9, or 10 and 11 cannot be distinguished anymore, and end up dividing their contributions in roughly equal shares, with residuals spread over other components. The spread is much smaller in the synthesis of components 1, 4, 5 and 7, which represent extreme metallicities within age groups. For intermediate Z systems (models B2, B3 and B6), however, these extreme components attract much of the percentage light contribution spilled over from neighboring populations of same age, so that their individual contributions in a real stellar population mix are as unreliable as the others.

This experiment demonstrates that for practical purposes the base is still linearly dependent, at least in a statistical sense, despite the effort of S91 in reducing the highly redundant original base of B88. This ‘noise-induced’, or ‘statistical linear dependence’, as we may call it, could indeed be inferred from the PCA analysis of S91, which showed that very little of the variance is associated with the last few eigenvectors. Within our probabilistic formulation, this ‘statistical linear dependence’ defines the structure of the likelihood function, which spreads first along directions in \mathbf{x} -space producing similar observables, carrying the mean \mathbf{x} to more densely populated regions. This effect explains the redistribution of the probability observed in Figs. 3, 4 and all other simulations.

It is therefore extremely difficult to retrieve accurately all the input stellar population parameters even for excellent data. This is further illustrated in Fig. 6, where the input parameters are plotted against their mean values, as derived from the Metropolis runs. One hundred artificial galaxies are plotted in each panel, with open circles corresponding to Set A observables and $S/N = 1000$, and crosses indicating the results for Set A and $S/N = 60$. A systematic underestimation of strong components is seen for $S/N = 60$, a good spectrum by any standard. This happens at the expense of an overestimation of the weaker components, which produces the large scatter seen in the bottom left corners of Fig. 6. The average uncertainties in the x_i ’s are of order $\sigma(x_i) \sim 3\text{--}4\%$ for $S/N = 60$. These are not enough to account for the large differences between input and output x_i ’s depicted in Fig. 6, so the errors induce a true bias in $\bar{\mathbf{x}}$. Note also that, in accordance with the discussion above, this bias is smaller for models with strong contributions of components 1, 4, 5 and 12, because of their extreme locations in the space of observables.

Only for unrealistically large S/N ’s, which far exceed the quality of the data used to build the base in the first place, one is able to break the ‘statistical linear dependence’ of the base by distinguishing fine details in the observables. This should *not* be interpreted as a failure of the method, as what the code actually synthesizes are the observables! These are *very precisely* reproduced by the *mean* (\mathbf{x} , \bar{A}_V), as illustrated in Fig. 7 for the W ’s of CaII K, CN, G-band and MgI and two continuum colors. The remaining observables, not shown in the figure, are equally well reproduced. Furthermore, whilst $S/N > 300$ is needed to recover accurately the model parameters from the mean solution, the agreement between the synthetic and measured observables is excellent for *any* S/N . Obviously, an even better agreement is obtained if instead of $(\bar{\mathbf{x}}, \bar{A}_V)$ the best model sampled during

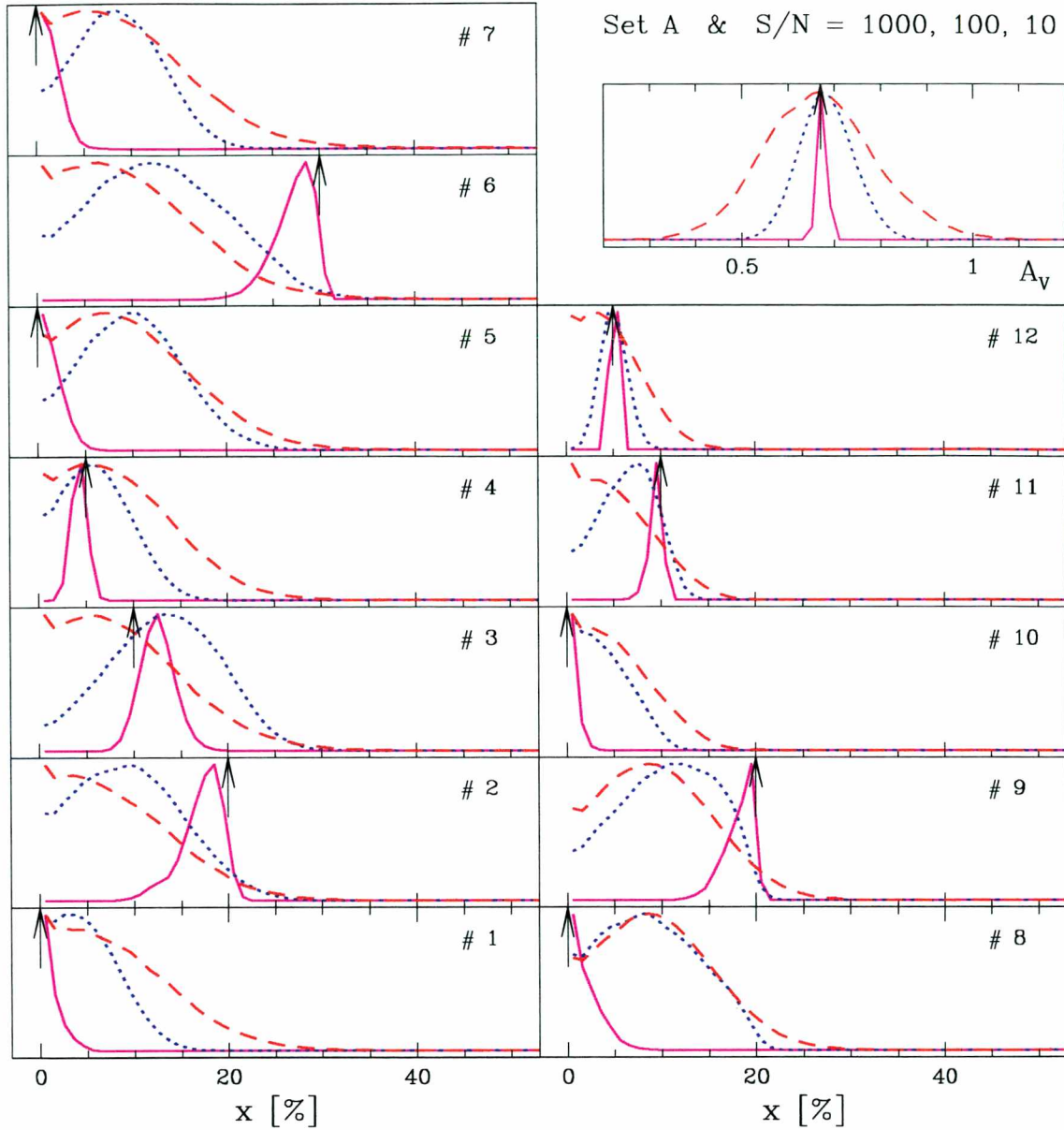


Figure 3. Effect of S/N upon the individual probability distributions of the synthesis parameters for Set A observables. The different lines correspond to $S/N = 1000$ (solid), 100 (dotted) and 10 (dashed). Notice how the probability of \mathbf{x} components with a large contribution (like x_2 , x_6 and x_9 in this example) are progressively redistributed among other components as the noise increases.

the Metropolis excursion is used to reconstruct the observables, as illustrated in the fourth column of plots in Fig. 7 (see §4.5).

In conclusion, this comparison of input and output observables shows that the spreading of the probability and the confusion between components seen in Figs. 3 and 5 is not an artifact of the method, but a consequence of the internal structure of the spectral base. As found in other EPS studies (O’Connell 1996 and references therein) synthesizing the observables is one thing, but trusting the detailed

star-formation history and chemical evolution implied by the synthesis is an altogether different story. On the positive side, a notable fact about Fig. 5 is that, consulting Table 1, one realizes that the reshuffling of the strength of the components occurs preferentially among populations of same *age*, an effect also clearly seen in Fig. 4. Though some redistribution among base elements of different age and *Z* in the older 1 and 10 Gyr bins due to the age-*Z* degeneracy also occurs, this indicates that the age distribution may be well recovered by the synthesis (§4.3).

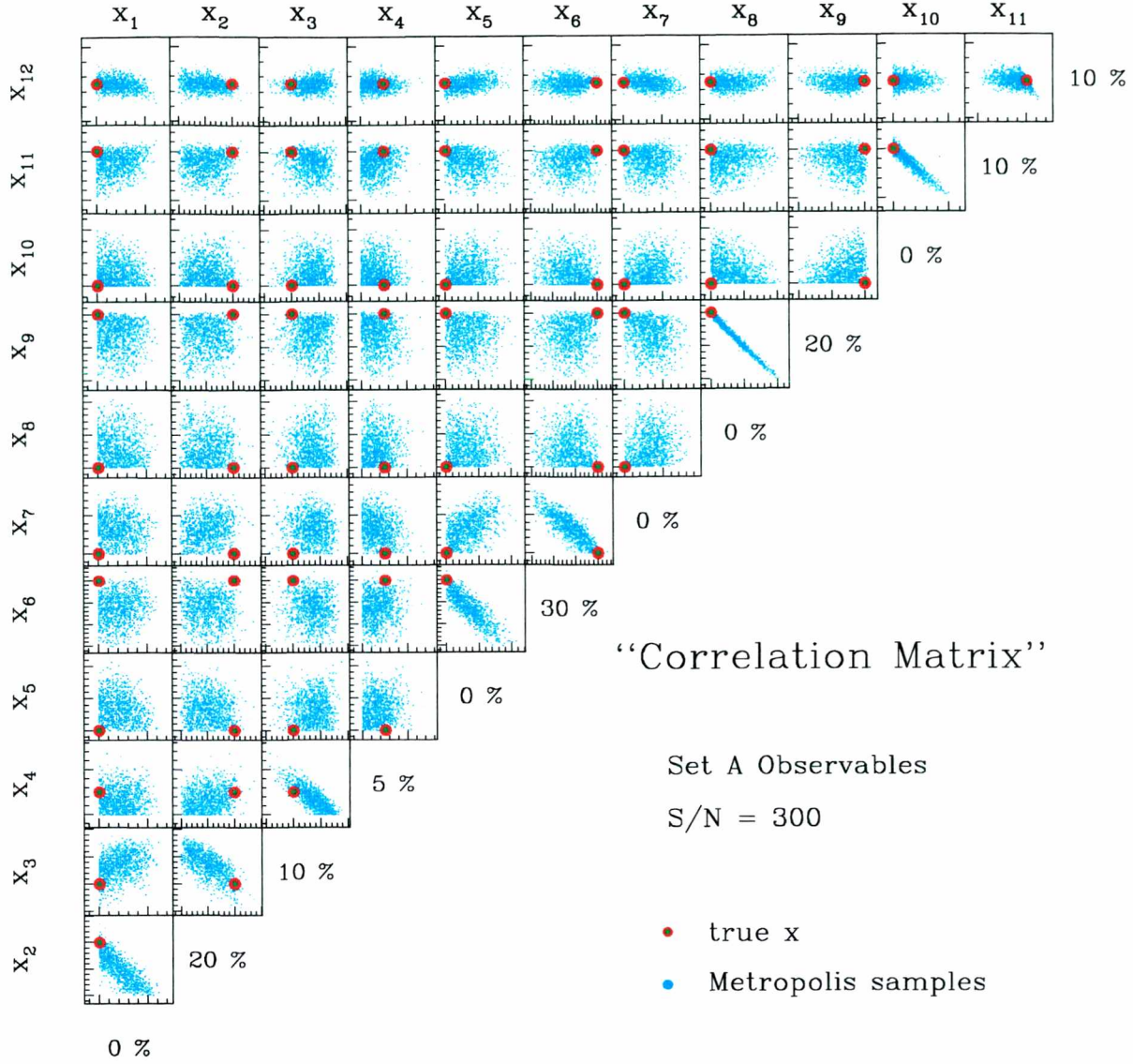


Figure 4. ‘Correlation Matrix’ for the test galaxy in Fig. 3. Each $x_i \times x_j$ panel shows one every 10000 of the first 10^7 steps of a Metropolis run for $(S/N)_{5870} = 300$ and Set A observables. The values at the bottom of each row indicate the input value of the corresponding component ($x_1 = 0$, $x_2 = 20\%$, \dots , $x_{11} = 10\%$, $x_{12} = 10\%$). These are also indicated by the solid circles in each panel. Small and big tick marks are spaced by 2 and 10% respectively in all plots. Note the strong anti-correlations between adjacent components of same age.

4.2 Effects of different sets of observables

The effects of which set of observables is used in the synthesis are illustrated in Fig. 8, where results for Sets A, B and C for the same test galaxy as in Fig. 3 and $S/N = 1000$ are compared. The figure shows that it is very advantageous to synthesize colors along with equivalent widths, despite the fact that one increases the dimensionality of the problem with the inclusion of A_V . This can be seen by comparing the performances of Sets A and B. The information contained

in the colors not only improves the estimation of \mathbf{x} but also allows a very good determination of A_V , which, unlike the population vector, is *always* well recovered by the synthesis.

Fig. 8 also shows how important it is to provide as many observational constraints as possible to constrain the synthesis, as Set A recovers the parameters much better than Sets B and C. Decreasing the number of observables in the synthesis thus have the same overall effect as decreasing the S/N . The width of the posteriors for Set C are partially attributed to the algebraic degeneracy in this set (8 observ-

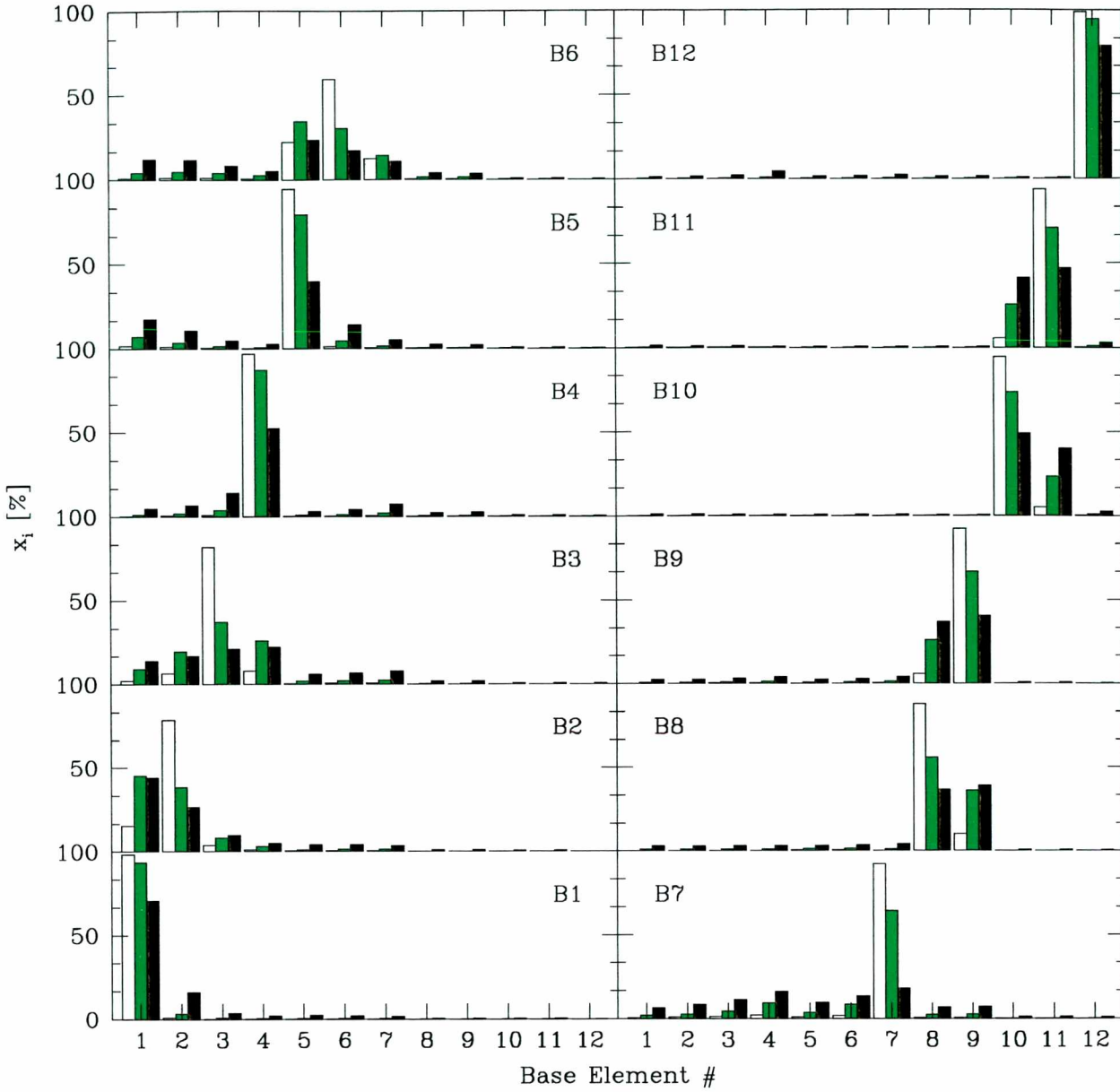


Figure 5. Synthesis of the base elements. Each panel shows the mean synthetic population vector (vertical axis) for 12 test galaxies whose 16 observables were generated with $x_i = 100\%$, $i = 1, \dots, 12$ (labeled B1...B12 respectively). Empty, shaded and filled histograms correspond to $S/N = 100, 30$ and 10 respectively. Ideally, all histograms should be concentrated in the component used to generate the observables. In practice, some bases elements, most notably 2, 3 and 6, are well synthesized by combinations of others even for large S/N .

ables and 12 degrees of freedom), which implies the existence of exact solutions in a sub-space of (\mathbf{x}, A_V) . That is not the case of Set A for the test galaxy in this example (Pelat 1998), but there, as for other sets, the ‘statistical linear dependence’ of the base is very efficient in spreading the likelihood even for such an idealized S/N ratio. The excellent agreement between the ‘observed’ and synthesized observables even for much worse S/N (Fig. 7) illustrates this point. Algebraic degeneracy is not as critical as the ‘statistical de-

generacy’ induced by the combined effects of noise, limited data and the internal correlations in the base.

Studies which synthesize only W ’s sometimes use the resulting population vector to compute a predicted continuum shape and derive A_V *a posteriori* through the comparison with the observed continuum (e.g., B88; Boisson et al. 2000). From our Set B simulations, we find that this procedure yields values of A_V 2–3 times less accurate than obtained with Set A. For $S/N = 60$, for instance, Set A yields

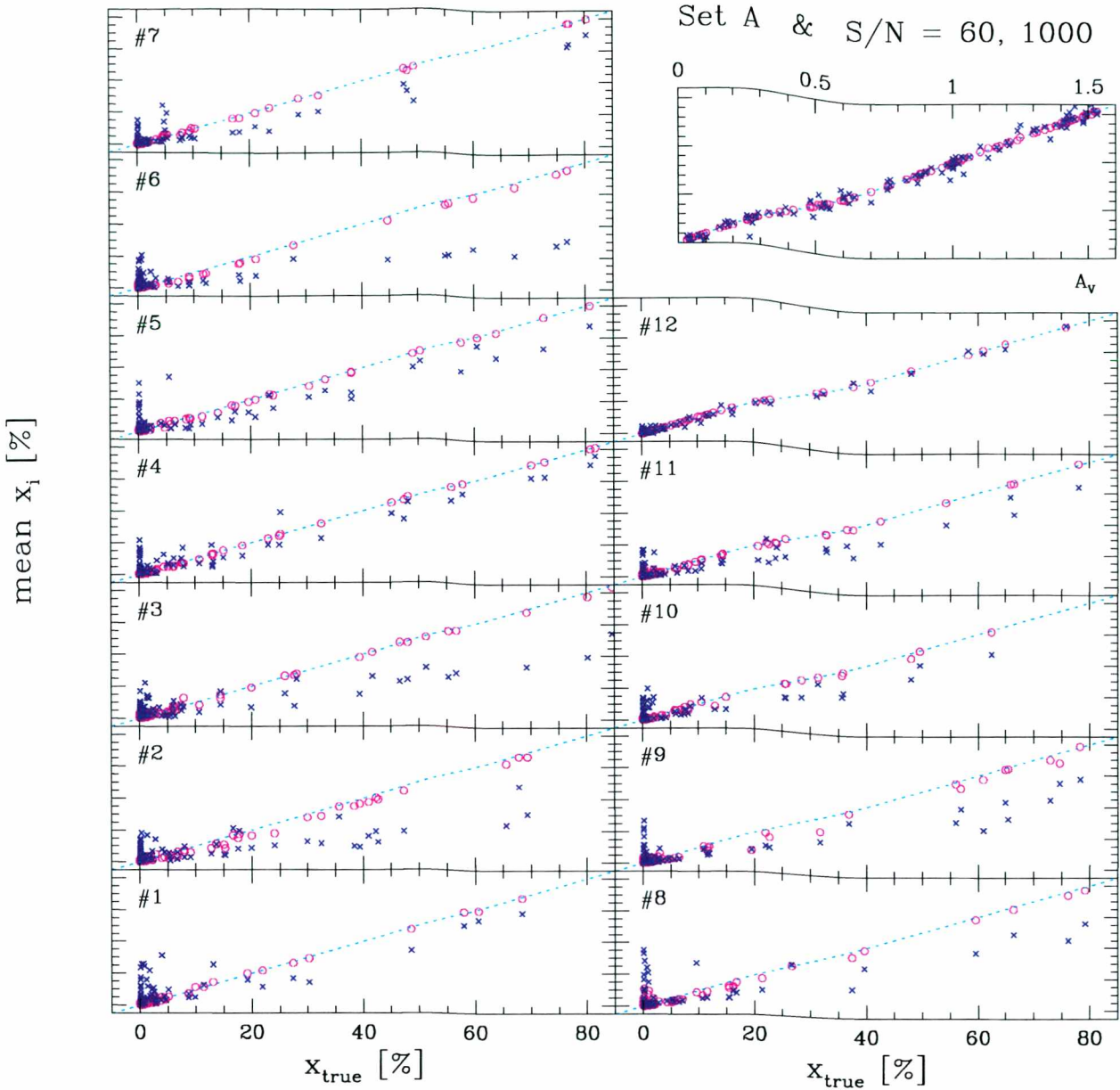


Figure 6. Input \times output plots for all 13 individual parameters in the synthesis. Open circles map the results for Set A and $S/N = 1000$, while crosses correspond to Set A and $S/N = 60$. The output parameters are the *mean* values mapped from the global likelihood function. The dotted, diagonal lines in each panel mark the ‘input = output’ line. Error bars on the model parameters are not shown for clarity. Notice the systematic underestimation of large x_i ’s and the corresponding overestimation of weak proportions for $S/N = 60$.

an rms dispersion in A_V of 0.06 mag around the true value, whereas the dispersion for set B is 0.13 mag. Since this is a relatively small difference, these experiments validate the *a posteriori* computation of the extinction. Colors are more useful to constrain the population vector than to estimate A_V .

4.3 Age grouped results

The results of the tests above reveal a striking difficulty to accurately determine all 12 base components of Bica’s base in the presence of even very modest noise and/or when not all base observables are available for the synthesis. As indicated by the simulations, both the measurement errors and the use of reduced sets of observables act primarily in the sense of spreading a strong contribution in one component preferentially among base elements of same age. *Group-*

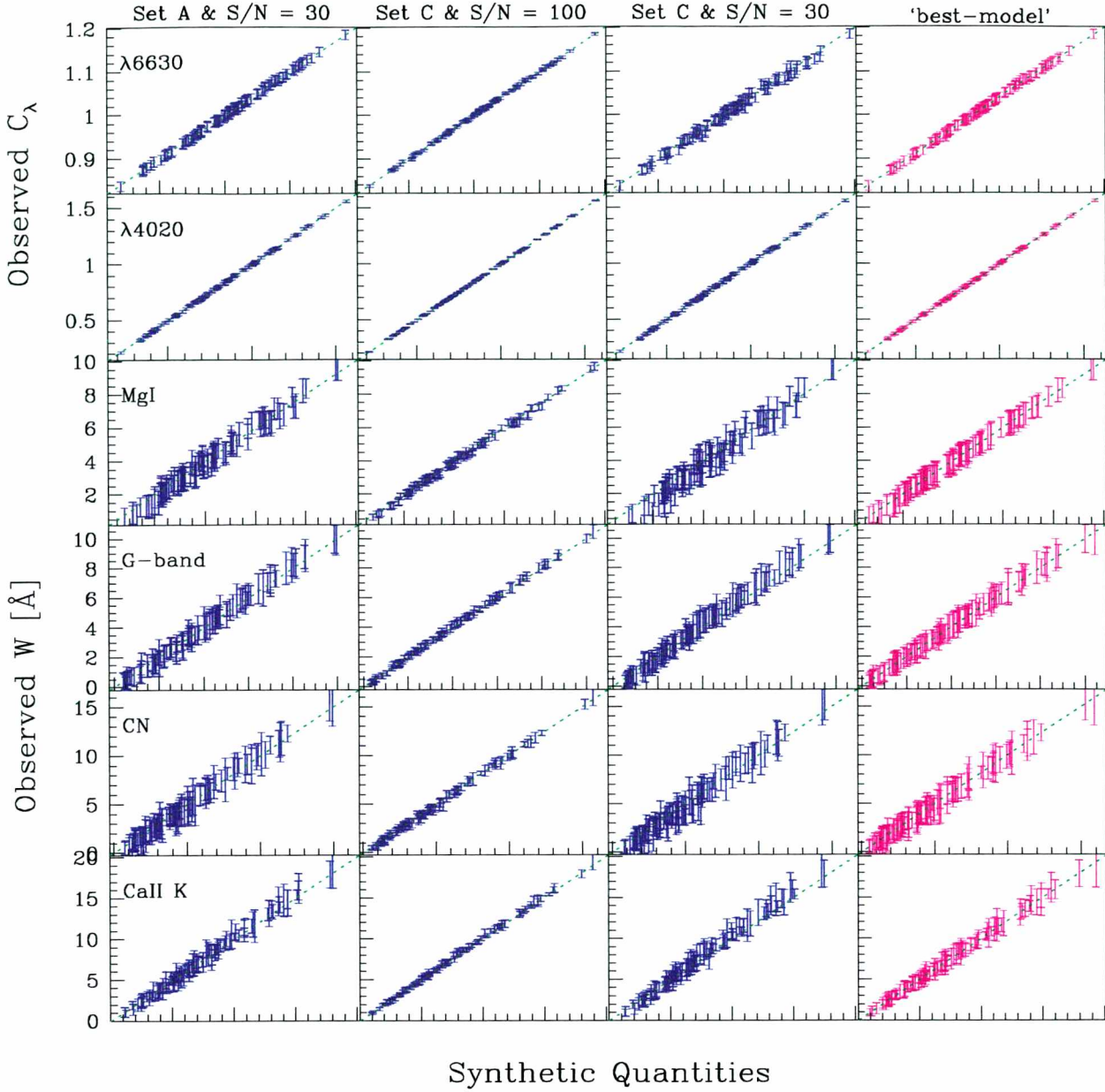


Figure 7. Input \times output observables for a series of 100 test-galaxies synthesized with the Metropolis-EPS sampler. The three left columns of plots correspond to the three indicated combinations of the set of observables and S/N . For these plots the x-axis quantities were computed with the *mean* (x, A_V) solution. The plots on the right column were made with Set C and $S/N = 30$ (as those in the third column), but the synthetic observables in this cases were computed with the best model found during the Metropolis runs. Only the equivalent widths of four absorption features and two colors exhibited, but results for the remaining observables are equally good. The vertical and horizontal scales are the same for each panel, and the input = output is indicated by a dotted line. Notice the excellent agreement, despite the fact that the input *parameters* are not accurately recovered by the synthesis process.

ing the population vector in age bins should thus produce more robust results. This expectation is confirmed in Fig. 9, where, analogous to what was done in Fig. 6, we plot the input values of the x_i 's against the output \bar{x}_i 's, but now for the five age binned groups, also for Set A simulations. The left panels correspond to the same values of S/N used in Fig. 6, so that one can appreciate the enormous improve-

ment achieved by grouping equal age elements. The right panels show that good results are also obtained for S/N of 30 (circles), and that young populations (≤ 100 Myr) are reasonably well traced even for S/N as low as 10 (crosses). Though some deviations survive, specially for the smaller S/N 's and the older age bins, the age-binned synthesis results are much more reliable than the results for the indi-

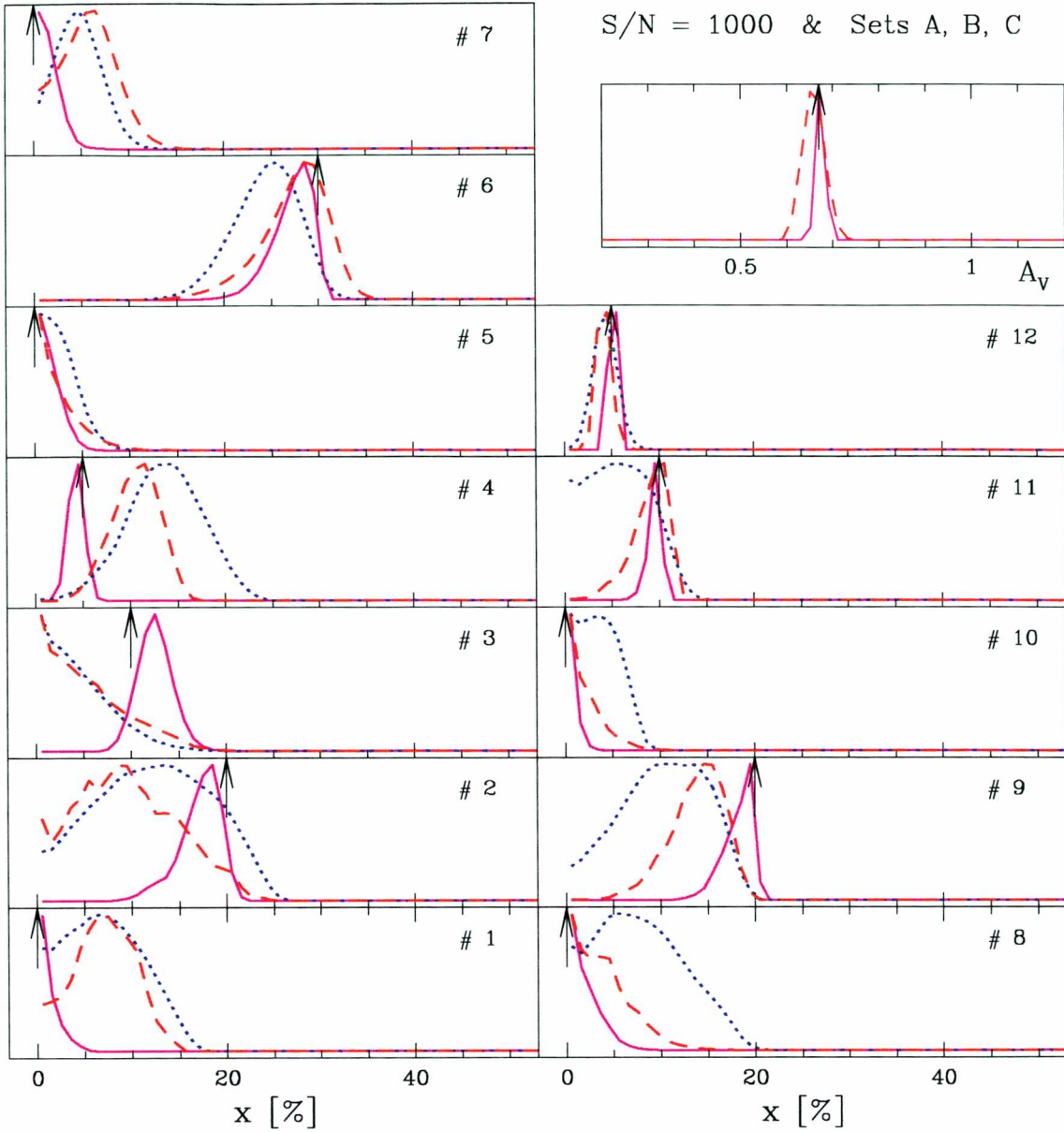


Figure 8. Analogous to Fig. 3, but exploring the effect of different sets of observables upon the synthesis parameters. Solid lines correspond to Set A (all observables), dotted lines to Set B (only 9 equivalent widths) and dashed lines to Set C (4 lines and 4 colors).

vidual components, and, most importantly, do *not* require unrealistic S/N .

This conclusion also holds for Sets B and C, as illustrated in Fig. 10. Naturally, the restricted input data in these sets translates into a loss of information, and the agreement is not as good as for Set A. Note that Set C does a much better job in recovering populations of 100 Myr or less than Set B, a further example of how useful colors are in the synthesis. This happens because the color information (present in Set C but absent in B) impose stronger constraints on young populations than in older systems.

The reliability of age-grouped proportions contrasts with the badly constrained and biased results for the individual components of \mathbf{x} , and is one of the main results of our study. This result puts in perspective all previous interpretations of the synthesis with this base. This fundamental limitation of the base was in fact previously known, but was never studied in detail. Schmitt et al. (1999), for instance, preferred to summarize the results of their synthesis study in terms of age grouped proportions for populations younger than 1 Gyr, in consonance with the results above. Still, in that work we kept a distinction between the different Z 's in

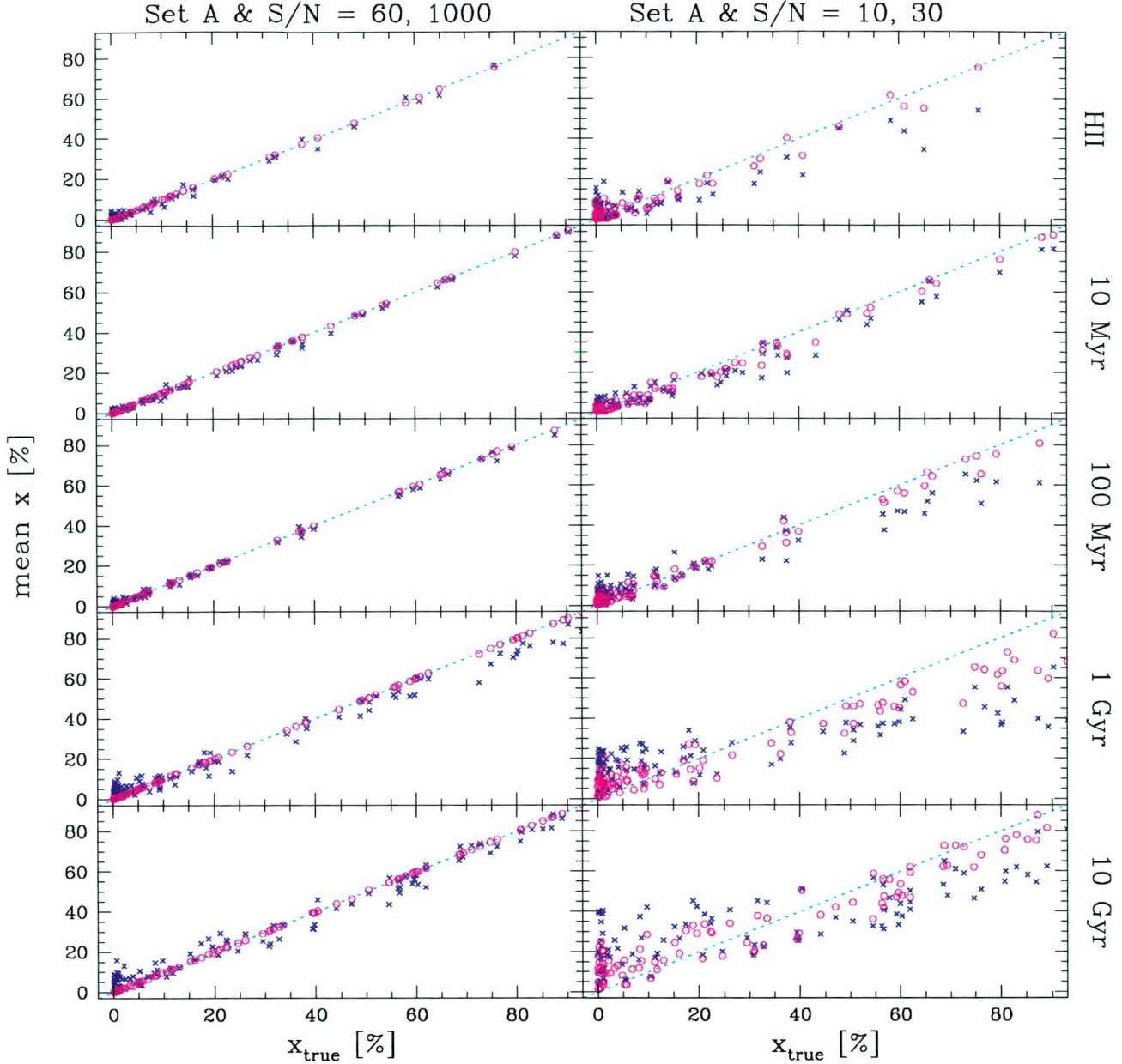


Figure 9. Age-binned light fractions in 100 input test galaxies versus the corresponding synthetic means. *Left panels:* Circles correspond to $S/N = 60$, and crosses to $S/N = 1000$, as in Fig. 6. *Right panels:* Circles correspond to $S/N = 30$, and crosses to $S/N = 10$. All results are for Set A observables.

the 10 Gyr age-group. At this level of detail the population fractions are very uncertain. In fact, since Set C observables and $S/N \sim 30\text{--}60$ were used in that study, Fig. 10 indicates that it would be safer to group their 1 and 10 Gyr proportions. Yet, the proportions for the younger (≤ 100 Myr) populations found by Schmitt et al. (1999), and which constituted the focus of that paper, are trustworthy (Fig. 10).

4.3.1 The age \times metallicity degeneracy

A close inspection of the $S/N = 10$ and 30 results in Fig. 9 reveals a compensation effect between the 1 and 10 Gyr age bins, albeit at a much smaller level than for the individual components (Fig. 6). This effect is much more pronounced in Sets B and C, as seen in Fig. 10.

The confusion between the 1 and 10 Gyr populations is related to the age- Z degeneracy, which sets in precisely in this age range. This well known effect (O’Connell 1986, 1994; Worthey 1994) is also present in Bica’s base and affects

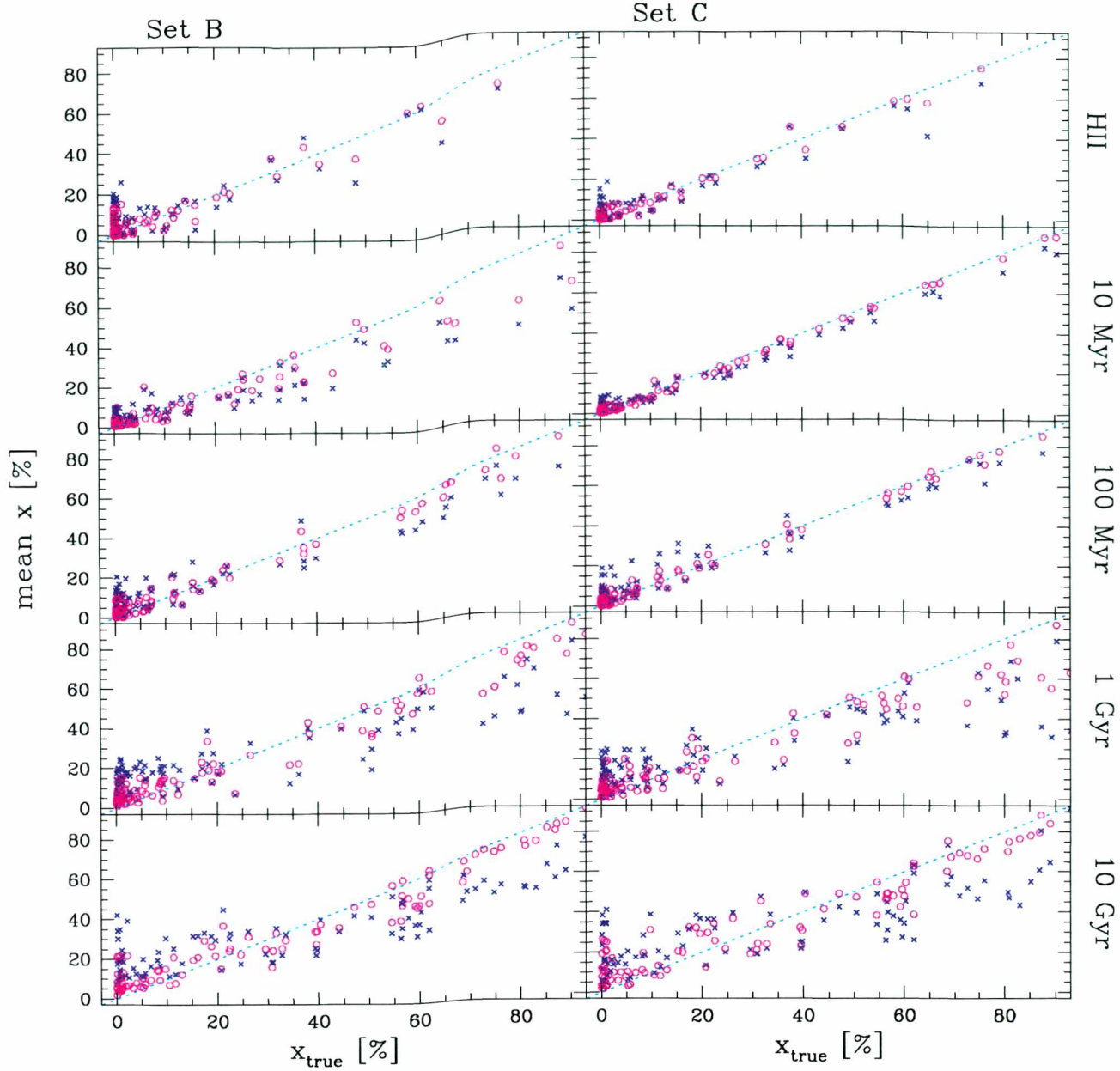


Figure 10. Input \times output age binned x fractions for Set B (left) and Set C (right) observables. Circles correspond to $S/N = 100$, and crosses to $S/N = 30$.

the synthesis results. To visualize and quantify this effect we have computed the average $\log Z$ and \log age for both the input and synthetic x for our test galaxies. Given that x is a light fraction at 5870 \AA , these averages ultimately represent flux-weighted ages and metallicities, which serve our current illustrative purposes, despite their ambiguous physical significance.

The ‘output - input’ \log age and $\log Z$ residuals are plotted against each other in Fig. 11. In these diagrams $\Delta \log \text{age} > 0$ (overestimated age) implies redder colors and stronger metal lines, which tend to be counter-balanced by

the bluer colors and weaker lines resulting from $\Delta \log Z < 0$ (underestimated Z). The age- Z degeneracy is most obvious in the results for Sets B and C, for which the biases in age and Z persist even for $S/N = 100$. For Set A, on the other hand, the residuals are small: $\lesssim 0.2$ dex for $S/N \gtrsim 60$, which shows that the age- Z degeneracy can be broken, at coarse the level of age and Z resolution offered by the base, with enough information and good spectra. Note that this remark applies to the global averages defined above, not the detailed component by component ages and Z ’s, which we concluded to be highly uncertain (Fig. 6).

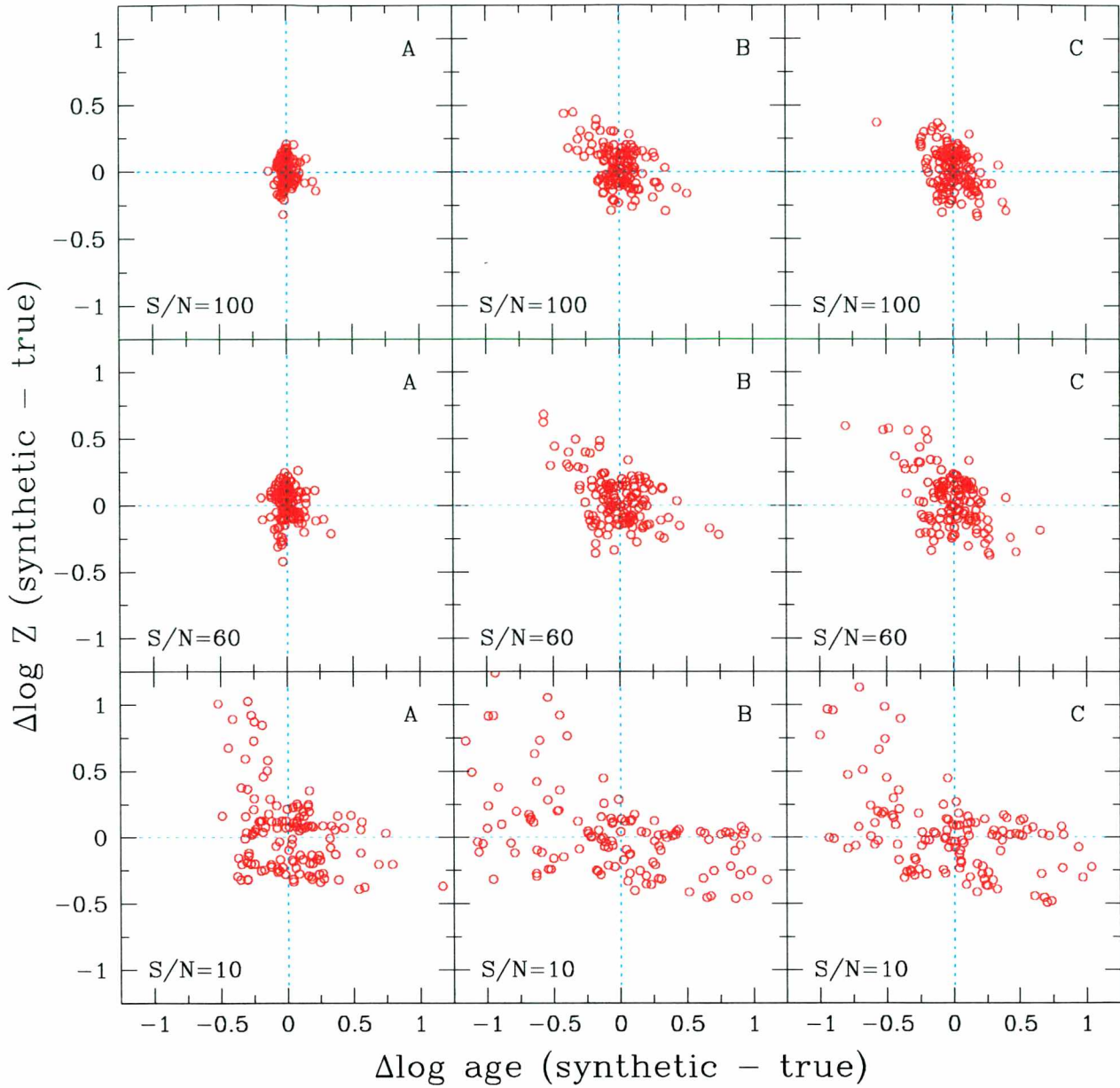


Figure 11. The age-metallicity degeneracy. The x-axis maps the difference between the output and input log age, with the corresponding difference in log Z plotted along the y-axis. Circles correspond to age and Z computed from the mean (\bar{x} , A_V), whereas dots correspond to the best model sampled. All 126 test galaxies are plotted. The different panels correspond to combinations of the data quality (S/N , labeled in the bottom-left corners) and the set of observables (top-right) used in the synthesis.

4.4 Metallicity grouped proportions

The results for Z binned proportions, plotted in Fig. 12, are not nearly as good as those for the age binned groups. With the exception of the $Z = 10^{-2} Z_{\odot}$ bin, which is actually represented by just one element (Table 1), the scatter in the input \times output Z -binned proportions is so large for all other Z 's that one is forced to conclude that no accurate description of the chemical history of galaxies can be afforded with this spectral base. Given this limited Z 'res-

olution power', one might consider removing intermediate components such as 2, 3 and 6. However, as we concluded that only the age distribution can be assessed with the base, there is no obvious advantage in this further reduction. This situation can be improved by providing *a priori* constraints on the occupancy of the age- Z plane, like those imposed by B88 based on chemical evolution arguments, since these effectively reduce the dimensionality of the base to typically 8 components, thus alleviating the confusion between components and producing better focused results. The validity of

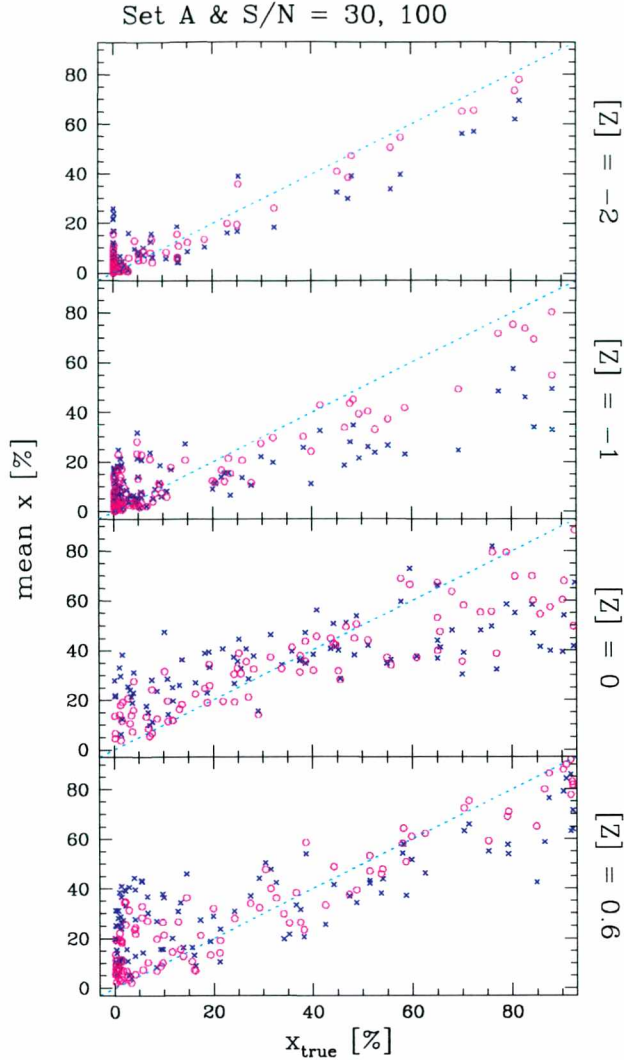


Figure 12. Metallicity binned light fractions in 100 input test galaxies versus the corresponding synthetic means. Circles correspond to $S/N = 100$, and crosses to $S/N = 30$, both for Set A observables. The ‘resolution power’ of the base is much better for ages (Fig. 9) than for metallicities.

this prior input external to the synthesis process, nonetheless, has to be evaluated in an object by object basis. In this study we followed S91 in not imposing any such extra information, as this approach encompasses a larger class of evolutionary scenarios and thus contemplates more possible applications.

4.5 Best \times Mean Parameters

Throughout these experiments, we have consistently used the *mean* parameters as an estimate of the result of the synthesis. This convention was deliberately chosen to follow the more widely employed version of EPS with this base and thus to allow a reassessment of previous results. Furthermore, the mean is a convenient and reasonable way to summarize the results insofar as it represents a ‘center of

mass’ of acceptable solutions. The severe biases identified in the mean synthetic population vector, however, raise the question of whether alternative approaches should be pursued instead.

By definition, minimization oriented techniques would do a better job at recovering the input parameters. Indeed, we verified that the best model population vector sampled in the Markov chain is much closer to \mathbf{x}_{true} than its expected value, an agreement which can only be improved given that the Metropolis sampler is not an optimal minimization procedure. Though the simulations confirmed our expectation that for small errors the mean and best solutions should both converge to the true parameters, it was surprising to realize how easily $\bar{\mathbf{x}}$ moves away from \mathbf{x}_{true} , and consequently also from \mathbf{x}_{best} . Very small uncertainties in the observables, at the levels of $S/N > 300$, are enough to set powerful compensation effects into operation (e.g., Figs. 3 and 4), shifting the components of $\bar{\mathbf{x}}$ away from their true values due to the asymmetric redistribution of the probability in \mathbf{x} -space.

Minimization procedures can in principle overcome this bias, as illustrated by the success of the S91 and Pelat (1997) tests as well as by our (not optimized) results for \mathbf{x}_{best} . A further bias in mean solutions is that they always produce $\bar{x}_i > 0$. Real galaxies, even if synthesizable by the base, are likely to have observables outside synthetic domain due to measurement errors. As shown by Pelat (1997, 1998), the best model in this case lies at the surface of the synthetic domain, where very often several of the \mathbf{x} -components are 0. In these aspects, minimization algorithms are more suitable than the more traditional sampling approach. However, the high susceptibility of the synthesis to the data quality indicates that minor perturbations in the input observables, within their errors, may seriously affect the estimation of the best model parameters. The stability of the best solution is thus likely to be more fragile than that of the mean solution, which takes into account the effects of the errors in the observables in the trade-off between an optimal and a statistically acceptable fit of the data. We have in fact detected this effect in a series of simulations, but a more careful study, with specialized minimization techniques is required to quantify it properly. It would be particularly interesting to carry out this test with the method developed by Pelat (1997) applied to the base used in this work. His Monte Carlo tests with a $n_{\star} = 10$ stellar base and $n_W = 19$ equivalent widths produced essentially unbiased population fractions even for S/N as low as 10. Extending his formalism to synthesize galaxy colors would also be desirable.

In any case, because of the nature of the EPS problem and the structure of Bica’s base, relatively large discrepancies in the parameters do not necessarily translate into large differences in the synthesized quantities. In this respect, the difference between mean and best models is largely academic, as the mean solution already does an excellent job in fitting the observables (Fig. 7).

5 SUMMARY

In this paper we have revisited the method of Empirical Population Synthesis with the aims of improving it both at the formal and computational levels, and exploring its

efficacy as a tool to probe the population mixture of galaxies. Our results can be divided into three parts.

(1) A simple probabilistic formulation of the problem was presented, which puts this method, traditionally employed in a less formal way, onto a mathematical footing. It was shown how former applications of the method fit into the probabilistic formulation, thus providing some *a posteriori* justification for previous results.

(2) An importance sampling scheme, based on the Metropolis algorithm, was developed and tested. It provides a more efficient and smooth mapping of the probability distribution of the parameters than is possible with the commonly used uniform grid sampling.

(3) In the third part of the paper, we applied the formulation and sampling method to a series of test galaxies constructed out the $n_* = 12$ star-clusters base of S91. The tests explored the ability to reconstruct the model parameters, which carry a record of the star formation and chemical evolution in galaxies, in the presence of (i) observational errors and (ii) limited data. This study centered on the comparison of input parameters with the *mean* synthetic solution (\bar{x}, \bar{A}_V), as this is the most common form of EPS in the literature with this base, and a systematic evaluation of the consistency of this method has not been carried out before. The main results of this study can be summarized as follows:

(a) The allowance for errors in the observables sparkle linear dependences within the base elements. This induces systematic *biases* in \bar{x} , redistributing the probability in components with a large fractional contribution to the integrated spectrum among similar components. This happens preferentially for components of same age. Though this was a predictable result, it was surprising to realize how powerful this effect is. For any realistic S/N , the bias is such that in most cases one cannot trust the estimates of all 12 *individual* components of \bar{x} to any useful level of confidence.

(b) Reducing the number of observational constraints, due to incomplete coverage of the 3300–9000 Å spectral interval spanned by the base, has the same overall effect as reducing the data quality, that is, to broaden, shift and skew the probability distribution of the x_i 's. We find that synthesizing colors as well as equivalent widths yields substantially better results than synthesizing W 's only. The need to account for the extinction as an extra parameter is largely compensated by the better focused x obtained by considering the color information, particularly for populations of 100 Myr and younger. Unlike for the individual population fractions, we do not detect systematic biases in the estimates of A_V .

(c) Despite the difficulty to retrieve accurately the detailed population vector, the observables are very precisely reproduced by the mean solution. The mapping between the parameters and observables spaces is thus highly 'degenerate'. The nature of this 'non-uniqueness' is not algebraic, as it happens also when the number of free parameters exceeds the numbers of observables and in undetermined cases where the solution is known to be unique. Instead, it is related to the possibility of mimicking to a statistically indistinguishable level the effects of certain base elements by linear combinations of other elements. This happens even for tiny

observational errors, inducing a 'spill over' of the likelihood among base components, carrying the mean solution along.

(d) The structure of the base reflects the fact that evolution is the main source of variance in stellar populations. Accordingly, we found that grouping the \bar{x} components in their five age bins produces much more reliable results, and we recommend this procedure when evaluating the detailed population vector obtained in previously published results with this base. Some confusion between the 1 and 10 Gyr populations also takes place due to the age- Z degeneracy, but this can be minimized with an adequate spectral coverage (measuring all observables in the base) and $S/N \gtrsim 30$. Grouping the results for the four different Z 's in the base is also recommended, but caution must be exerted when interpreting the results since these are much less precise than the age-binned proportions.

The difficulties to retrieve accurately the true stellar population parameters stem from two sources: (i) The convention to represent the results of the synthesis by the mean solution, and (ii) limitations imposed noise-induced linear dependences in the base. Minimization techniques may be instrumental in overcoming the biases in the mean population vector, which ultimately limit the resolution of the 12-D base to its 5 age bins. Within our probabilistic formulation, other priors can be investigated. Maximum Entropy, for example, is a prior that has been applied successfully to inverse problems formally similar to population synthesis, like image restoration. However, more than improvements on the method, which are possible and desirable, further work on the base itself will certainly be needed. Indeed, the limited age and Z 'resolution power' of the base employed here lies at the heart of all difficulties identified in this paper. Including other spectral indices, modeling the full spectrum and incorporating other relevant information such as mass to light ratios in the synthesis process are possible ways to progress in this direction.

ACKNOWLEDGEMENTS

We thank Catherine Boisson, Monique Joly, Jihane Moulata and Didier Pelat for stimulating discussions on the virtues and drawbacks of EPS techniques. RCF thanks the hospitality of Johns Hopkins University, where this work was finalized. Support for this work was provided by the National Science Foundation through grant # GF-1001-99 from the Association of Universities for Research in Astronomy, Inc., under NSF cooperative agreement AST-9613615. JRSL acknowledges an MSc fellowship awarded by CAPES. Partial support from CNPq, FAPESP, PRONEX and FUNPESQUISA-UFSC is also acknowledged.

REFERENCES

- Bica E., 1988, A&A, 195, 76
- Bica E., Alloin D., 1986a, A&A, 162, 21
- Bica E., Alloin D., 1986b, A&AS, 66, 171
- Bica E., Alloin D., 1987, A&A, 186, 49
- Bica E., Alloin D., Schmidt A. A., 1990, A&A, 228, 23
- Bica E., Alloin D., Schmitt H. R., 1994, A&A, 283, 805
- Bonatto C., Bica E., Alloin D., 1995, A&AS, 112, 71

- Boisson C., Joly M., Moulata J., Pelat D., Serote Roos M., 2000, *A&A*, 357, 850
- Bonatto C., Bica E., Pastoriza M. G., Alloin D., 1996, *A&AS*, 118, 89
- Bonatto C., Pastoriza M. G., Alloin D., Bica E., 1998, *A&A*, 334, 439
- Bonatto C., Bica E., Pastoriza M. G., Alloin D., 1999, *A&A*, 343, 100
- Bonatto C., Bica E., Pastoriza M. G., Alloin D., 2000, *A&A*, 355, 99
- Bruzual G., Charlot S., 1993, *ApJ*, 405, 538
- Charlot S., Bruzual G., 1991, *ApJ*, 367, 126
- Cardelli J. A., Clayton G. C., Mathis, J. S. 1989, *ApJ*, 345, 245
- Cid Fernandes R., Storch-Bergmann T., Schmitt H. R., 1998, *MNRAS*, 297, 579
- De Mello D. F., Keel W. C., Sulentic J. W., Rampazzo R., Bica E., White R. E. III, 1995, *A&A*, 297, 331
- Faber S. M., 1972, *A&A*, 20, 361F
- Fioc M., Rocca-Volmerange B., 1997, *A&A*, 326, 950
- Guiderdoni B., Rocca-Volmerange B., 1987, *A&A*, 186, 1
- Jablonka P., Alloin D., 1995, *A&A*, 298, 361
- Jablonka P., Alloin D., Bica E., 1990, *A&A*, 235, 22
- Joly M., 1974, *A&A*, 33, 177
- Kong X., Cheng F. Z., 1999, *A&A*, 351, 477
- Larson R. B., Tinsley B. M., 1974, *ApJ*, 192, 293
- Larson R. B., Tinsley B. M., 1978, *ApJ*, 219, 46
- Leão J. R. S., 2001, MSc Thesis, Universidade Federal de Santa Catarina
- Leitherer C. et al., 1996, *PASP*, 108, 996L
- Leitherer C., Schaerer D., Goldader, J. D., Delgado, R. M., González R., Carmelle K., Denis F., de Mello D. F.; Devost D., Heckman T. M., 1999, *ApJS*, 123, 3
- MacKay D., 2001, Information Theory, Pattern Recognition and Neural Networks, <http://wol.ra.phy.cam.ac.uk/mackay/>
- Metropolis N., Rosenbluth A., Teller A., Teller E., 1953, *J. of Chem. Phys.*, 21, 1087.
- Moulata J., Pelat D., 2000, *MNRAS*, 314, 409
- O'Connell R. W., 1976, *ApJ*, 206, 370
- O'Connell R. W., 1986, in *Stellar Populations*, ed. C. Norman, A. Renzini, M. Tosi (Cambridge: Cambridge Univ. Press), p.213
- O'Connell R. W., 1994, in *Nuclei of Normal Galaxies: Lessons from the Galactic Center*, ed. R. Genzel (Dordrecht: Kluwer),
- O'Connell R. W., 1996, in *From Stars to Galaxies: The impact of stellar physics on galaxy evolution*, eds. C. Leitherer, U. Fritze-von-Alvensleben, J. Huchra (ASP: San Francisco),
- Pelat D., 1997, *MNRAS*, 284, 365
- Pelat D., 1998, *MNRAS*, 299, 877
- Pickles A. J., 1985, *ApJ*, 296, 340
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, *Numerical Recipes*, 2nd. edition (Cambridge: Cambridge University Press)
- Raimann D., Bica E., Storch-Bergmann T., Melnick J., Schmitt H. R., 2000, *MNRAS*, 314, 295
- Serote Roos M., Boisson C., Joly M., Ward M., 1998, *MNRAS*, 301, 1
- Serote Roos M., Boisson C., Joly M. 1996, *A&A*, 117, 93
- Schmidt A. A., Copetti M. V. F., Alloin D., Jablonka P., 1991, *MNRAS*, 249, 766
- Schmidt A. A., Bica E., Alloin D., 1990, *MNRAS*, 243, 620
- Schmidt A. A., Bica E., Dottori H. A., 1989, *MNRAS*, 238, 925
- Schmitt H. R., Bica E., Pastoriza, M. G., 1996, *MNRAS*, 278, 965
- Schmitt H. R., Storch-Bergmann T., Cid Fernandes R., 1999, *MNRAS*, 303, 173
- Silva D., Cornell M., 1992, *ApJS*, 81, 865
- Spinrad H., Taylor B. J., 1971, *ApJS*, 22, 445
- Tinsley B. M., 1972, *A&A*, 20, 383
- Worthey G., 1994, *ApJS*, 95, 107

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Aloisi A., Leitherer C., 2000, AAS, 197, 3001A.
- [2] Barbuy B., Renzini A., 1991, IAU Symp. 149: The Stellar Populations of Galaxies, 149.
- [3] Bica E., 1988, A&A, 195, 76.
- [4] Bica E., Alloin D., 1986a, A&A, 162, 21.
- [5] Bica E., Alloin D., 1987, A&A, 186, 49.
- [6] Bica E., Alloin D., Schmidt A. A., 1990, A&A, 228, 23.
- [7] Bica E., Alloin D., Schmitt H. R., 1994, A&A, 283, 805.
- [8] Boisson C., Joly M., Moulata J., Pelat D., Serote Roos M., 2000, A&A, 357, 850
- [9] Bonatto C., Bica E., Pastoriza M. G., Alloin D., 1996, A&AS, 118, 89.
- [10] Bonatto C., Pastoriza M. G., Alloin D., Bica E., 1998, A&A, 334, 439.
- [11] Bonatto C., Bica E., Pastoriza M. G., Alloin D., 1999, A&A, 343, 100.
- [12] Bonatto C., Bica E., Pastoriza M. G., Alloin D., 2000, A&A, 355, 99.
- [13] Bruzual G., Charlot S., 1993, ApJ, 405, 538.
- [14] Cardelli J. A., Clayton G. C., Mathis, J. S. 1989, ApJ. 345, 245.
- [15] Chiosi C., Renzini A., 1986, Book-Review - Spectral Evolution of Galaxies, 233, 993C.
- [16] Cid Fernandes R., Storch-Bergmann T., Schmitt H. R., 1998, MNRAS, 297, 579.
- [17] Cid Fernandes R., Sodré L., Schmitt H. R., Leão J. R. S., 2001, MNRAS, 325, 60.
- [18] Levenson N. A., Cid Fernandes R., Weaver K. A., Heckman T. M., Storch-Bergmann T., 2001b, ApJ, 557, 54.

- [19] De Mello D. F., Keel W. C., Sulentic J. W., Rampazzo R., Bica E., White R. E. III, 1995, A&A, 297, 331.
- [20] Faber S. M., 1972, A&A, 20, 361F.
- [21] Folks S., 1999, MNRAS, 308, 459.
- [22] Fioc M., Rocca-Volmerange B., 1997, A&A, 326, 950.
- [23] Jablonka P., Alloin D., 1995, A&A, 298, 361.
- [24] Kong X., Cheng F. Z., 1999, A&A, 351, 477.
- [25] Kennicutt R. C. Jr., 1992, ApJS, 79, 255.
- [26] Larson R. B., Tinsley B. M., 1974, ApJ, 192, 293.
- [27] Larson R. B., Tinsley B. M., 1978, ApJ, 219, 46.
- [28] Leitherer C. *et al.* , 1996, PASP, 108, 996L.
- [29] Leitherer C., Schaerer D., Goldader, J. D., Delgado, R. M., González R., Carmelle K., Denis F., de Mello D. F.; Devost D., Heckman T. M., 1999, ApJS, 123, 3.
- [30] Leitherer C., Schaerer D., Goldader, J. D., Delgado, R. M., González R., Carmelle K., Denis F., de Mello D. F.; Devost D., Heckman T. M., 1999, ApJS, 123, 3.
- [31] Leitherer C., Leão J. R. S., Heckman T.M., Lennon D. J., Pettini M., Robert C., 2000, ApJ, 550, 724.
- [32] MacKay D., 2001, Information Theory, Pattern Recognition and Neural Networks, <http://wol.ra.phy.cam.ac.uk/mackay/>.
- [33] Metropolis N., Rosenbluth A., Teller A., Teller E., 1953, J. of Chem. Phys., 21, 1087.
- [34] Moore E., 1968 PhD Dissertation, University of arizona.
- [35] Moulata J., Pelat D., 2000, MNRAS, 314, 409.
- [36] O'Connell R. W., 1976, ApJ, 206, 370.
- [37] O'Connell R. W., 1986, in Stellar Populations, ed. C. Norman, A. Renzini, M. Tosi (Cambridge: Cambridge Univ. Press), p.213.
- [38] O'Connell R. W., 1994, in Nuclei of Normal Galaxies: Lessons from the Galactic Center, ed. R. Genzel (Dordrecht: Kluwer)
- [39] O'Connell R. W., 1996, in From Stars to Galaxies: The impact of stellar physics on galaxy evolution, eds. C. Leitherer, U. Fritze-von-Alvensleben, J. Huchra (ASP: San Francisco).

- [40] Pelat D., 1997, MNRAS, 284, 365.
- [41] Pelat D., 1998, MNRAS, 299, 877.
- [42] Pickles A. J., 1985, ApJ, 296, 340.
- [43] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes, 2nd. edition (Cambridge: Cambridge University Press)
- [44] Raimann D., Bica E., Storchi-Bergmann T., Melnick J., Schmitt H. R., 2000, MNRAS, 314, 295.
- [45] Serote Roos M., Boisson C., Joly M., Ward M., 1998, MNRAS, 301, 1.
- [46] Serote Roos M., Boisson C., Joly M. 1996, A&A, 117, 93.
- [47] Schmidt A. A., 1989.
- [48] Schmidt A. A., Copetti M. V. F., Alloin D., Jablonka P., 1991, MNRAS, 249, 766.
- [49] Schmidt A. A., Bica E., Alloin D., 1990, MNRAS, 243, 620.
- [50] Schmidt A. A., Bica E., Dottori H. A., 1989, MNRAS, 238, 925.
- [51] Schmitt H. R., Storchi-Bergmann T., Cid Fernandes R., 1999, MNRAS, 303, 173.
- [52] Schmitt H. R., Bica E., Patoriza M. G., 1996, MNRAS, 278, 965.
- [53] Spinrad H., Taylor B. J., 1971, ApJS, 22, 445.
- [54] Tinsley B. M., 1972, A&A, 20, 383.
- [55] Worthey G., 1994, ApJS, 95, 107.
- [56] Wood D. B., 1996, ApJ, 145, 36.
- [57] York d. G., 2000, AJ, 120, 1579.